UNIVERSITY OF NATIONAL AND WORLD ECONOMY

The Spirit Makes The Power

3RD INTERNATIONAL CONFERENCE ON APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION

ICAICTSEE – 2013

December 6 – 7th, 2013

University of National and World Economy

Sofia, Bulgaria

CONFERENCE PROCEEDINGS

UNDER THE AUSPICES OF:

ifip

the leading edge of information technology

PUBLICATIONS INDEXED BY:

EBSCO HOST

and

ProQuest

SPONSORS AND PARTNERS:

dWare

MusalaSoft

кадинама

Microsoft

PROCEEDINGS OF THE

INTERNATIONAL CONFERENCE ON APPLICATION OF

INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION

ICAICTSEE-2013

EDITOR:

PROF. DR. DIMITER G. VELEV

DEPT. OF INFORMATION TECHNOLOGIES AND COMMUNICATIONS

UNIVERSITY OF NATIONAL AND WORLD ECONOMY

UNSS – STUDENTSKI GRAD

1700 SOFIA, BULGARIA

DGVELEV@UNWE.BG

ISSUED FOR PUBLICATION: JULY 24TH 2014

SIZE: JIS-B5

QUIRES: 49

VOLUME: 200

ISBN 978-954-644-586-5

CONFERENCE COMMITTEES

CONFERENCE CHAIR

Dimiter G. **Velev**	University of National and World Economy	***Bulgaria***

PROGRAM COMMITTEE

César Correa **Arias**	Universidad de Guadalajara	***Mexico***
Saji **Baby**	GEO Environmental Consultation	***Kuwait***
Victor **Blagodatskikh**	Moscow State University of Economics, Statistics and Informatics	***Russia***
Helena **Blazun**	University of Maribor	***Slovenia***
Robert **Brumnik**	Institute for Corporate Security Studies	***Slovenia***
Gaetano **Cascini**	Politecnico di Milano	***Italy***
Donald T. **Chang**	Metropolitan State College of Denver	***USA***
Sérgio Manuel Serra da **Cruz**	Universidade Federal Rural do Rio de Janeiro	***Brazil***
Tadeusz **Czachorski**	Institue of Theoretical and Applied Informatics, PAS	***Poland***
Vladimir **Dimitrov**	Sofia State University	***Bulgaria***
Ivan **Drogobytsky**	Fiancial University under the Goverment of the Russian Federation	***Russia***
Francisco Patricio **Esteves**	Higher Institute of Technology and Sciences	***Angola***
Vasil **Georgiev**	Sofia State University	***Bulgaria***
Valentin **Goev**	University of National and World Economy	***Bulgaria***
Bernard **Grabot**	University of Toulouse	***France***
Christopher Khoo Soo **Guan**	Nanyang Technological University	***Singapore***
James K. **Ho**	Iniversity of Illinois at Chicago	***USA***
Aytekin **Isman**	Sakarya University	***Turkey***
Sergey **Ivanov**	University of the District of Columbia, Washington, DC	***USA***
Masayasu **Kanno**	Kanagawa University	***Japan***
Madjid Malikovich **Karimov**	Tashkent University of Information Technologies	***Uzbekistan***
Sergii **Kavun**	Kharkiv Institute of Banking	***Ukraine***
Alexander **Khoroshilov**	UNESCO Institute for Information Technologies in Education	***Russia***
Leszek F. **Korzeniowski**	European Association for Security	***Poland***
Hee Dong **Kim**	Hankuk University of Foreign Studies	***S.Korea***
Valentin **Kisimov**	University of National and World Economy	***Bulgaria***

Peter **Kokol**	University of Maribor	***Slovenia***
George **Kovacs**	Computer and Automation Institute	***Hungary***
Oleksiy E. **Kovalenko**	Institute of Mathematical Machines and Systems Problems, NASU	***Ukraine***
Anatoly **Krapivensky**	Volgograd branch of Moscow University of Fiance and Law	***Russia***
Vinod **Kumar**	Sprott School of Business, Carlton University	***Canada***
Andrew **Kusiak**	The University of Iowa	***USA***
Ronald S. **Laura**	University of Newcastle	***Australia***
Patrick **Letouze**	Universidade Federal do Tocantins	***Brazil***
Havar Amir oglu **Mammadov**	Azerbaijan Technical University	***Azerbaijan***
Violeta **Manevska**	St. Kliment Ohridski University - Bitola	***Macedonia***
Christopher **McMahon**	University of Bristol	***UK***
Pedja M. **Milosavljevic**	University of Nis	***Serbia***
A.K. **Mishra**	Motilal Nehru National Institute of Technology	***India***
Igor **Mladenovic**	University of Nis	***Serbia***
Alexander N. **Moiseev**	Tomsk State University	***Russia***
Balakrishnan **Muniandy**	Universiti Sains Malaysia	***Malaysia***
Valentin V. **Nechaev**	Moscow State Institute of Radiotechnics, Electronics and Automatics	***Russia***
Tetsuo **Noda**	Shimane University	***Japan***
Serghei **Ohrimenco**	Laboratory of Information Security, AESM	***Moldova***
Gustav **Olling**	IFIP TC5	***USA***
K.R. **Pardasani**	Maulana Azad National Institute of Technology	***India***
Veselka **Pavlova**	University of National and World Economy	***Bulgaria***
David Nadler **Prata**	Universidade Federal do Tocantins	***Brazil***
A. M. **Rawani**	National Institute of Technology	***India***
Volodymyr **Shkir**	National University Lviv Polytechnic	***Ukraine***
Yixun **Shi**	Bloomsburg University of Pennsylvania	***USA***
Santhidran **Sinnappan**	University Tunku Adbul Rahman	***Malaysia***
Kamelia **Stefanova**	University of National and World Economy	***Bulgaria***
Karen **Starr**	Alfred Deakin Research Institute	***Australia***
Olexandr **Trydid**	Kharkiv Institute of Banking	***Ukraine***
A.F. **Verlan**	Pukhov Institute for Modelling in Energy Engineering, NASU	***Ukraine***
Hsin-Hung **Wu**	National Changhua University of Education	***Taiwan***
Shi **Yizhe**	ShenYang University of Chemical Technology	***China***
Parviz Ali **Zade**	OKAN University	***Turkey***
Milan **Zorman**	University of Maribor	***Slovenia***

LOCAL ORGANIZING COMMITTEE

Emil **Denchev**	University of National and World Economy	***Bulgaria***
Rosen **Kirilov**	University of National and World Economy	***Bulgaria***
Mihail **Konchev**	University of National and World Economy	***Bulgaria***
Vania **Lazarova**	University of National and World Economy	***Bulgaria***
Plamen **Milev**	University of National and World Economy	***Bulgaria***
Alexandrina **Murgeva**	University of National and World Economy	***Bulgaria***
Monika **Tzaneva**	University of National and World Economy	***Bulgaria***
Ilko **Velikov**	University of National and World Economy	***Bulgaria***

CONTENTS

Dear Colleagues,

On behalf of the Conference Committees I have the honor and pleasure to thank you all for your true participation in the work of the International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE-2013) which took place on December 6 – 7th, 2013 at the University of National and World Economy (UNWE), Sofia, Bulgaria.

I would like specially to express my gratitude to all of the program committee members for their genuine support without which it would never had happened.

Due to the current state of scientific development in all spheres of human activity, the constant knowledge and skill actualization of the academia and researchers in the field of Information and Communication Technology (ICT) is an obligatory necessity, especially when the world is in a global economic and financial crisis. The definition of long-term scientific research tasks in this area has a strategic importance, which is even truer for young researchers, lecturers and Ph.D. students.

The conference aims were conducting analysis of the current problems and presenting results of the ICT application in different areas of economy, education and related areas of scientific knowledge; outlining the existing possibilities for the application of modern ICT tools, methods, methodologies and information systems in economy and education; discussing advanced and emerging research trends with a long-term importance in the field of ICT application in economy and education.

The conference has established an academic environment that fosters the dialogue and exchange of ideas between different levels of academic and research communities.

The conference outcome is 107 published research papers (33 from foreign participants), the explosion of fresh ideas and the establishment of new professional relations.

I do hope to meet you all again at ICAICTSEE – 2014.

Dimiter G. Velev
ICAICTSEE-2013 Chair

Classical Web-Based Client-Server Systems and the ACID Principles of the Transactional Approach

Vanya Lazarova

University of National and World Economy,
e-mail:vlazarova@unwe.bg

Abstract. In this article are considered the four classical principles of transactional approach in client/server systems. We define as "Classical" client-server systems the ones that are developed specifically for one company. ACID (Atomicity, Concurrency, Independence and Durability) are the four classical principles which are discussed in the paper. Classical client-server systems we are considering as non-distributed and distributed. The four classical principles has applied in a different way according to the system - non-distributed or distributed.

Keywords: ACID principles, client/server system, non-distributed system, distributed system

1. Introduction

This article defines as "Classical" client-server systems that are developed specifically for one company. A classical client-server system could be implemented in a local network. In this case the computers are typically placed in the building of the company and managed by the employees from the company's administrative department. Another kind of classical client-server system is the web-based system, which are working on a set of computers connected via Internet.

There are two big groups of the web-based "Classical" client/server applications - non-distributed and distributed applications.

Non-distributed applications are one component applications with non-autonomous parts. The components are shared by the users all the time and all resources are accessible.

A distributed system consists of computers, connected through a network and distribution middleware, which enables computers to coordinate their activities and to share the resources, so that from the users point of view the system as one facility. The distributed applications can be the cloud based applications and client-server applications, but all non-distributed applications are only classical (see fig. 1).

	Classical client/server applications	Cloud applications
Non-distributed applications	+	-
Distributed applications	+	+

Figure 1. Classical non-distributed and distributed client/server applications

2. Classical web-based client-server systems and the ACID principles of the transactional approach

Almost any classical web-based client-server system includes the following three types of generalized functions:

- User interface - functions implementing the information system interaction with the user.
- Business logic - functions of information processing in accordance with specific rules governing the specific application area. Description of the business logic includes descriptions of business rules and business processes. Business rules are the operations, definitions and restrictive conditions applied by the organization in the achievement of its goals.
- Storage and retrieval of data from databases or from individual files. Databases can be different type – a relational or a hierarchical. The relational databases also can be any tape (Radoev, M., 2012).

When the functions of the information system are divided into software components so that they can be installed to run on a separate machine, they form a tier (layer) of information system (see fig. 2).

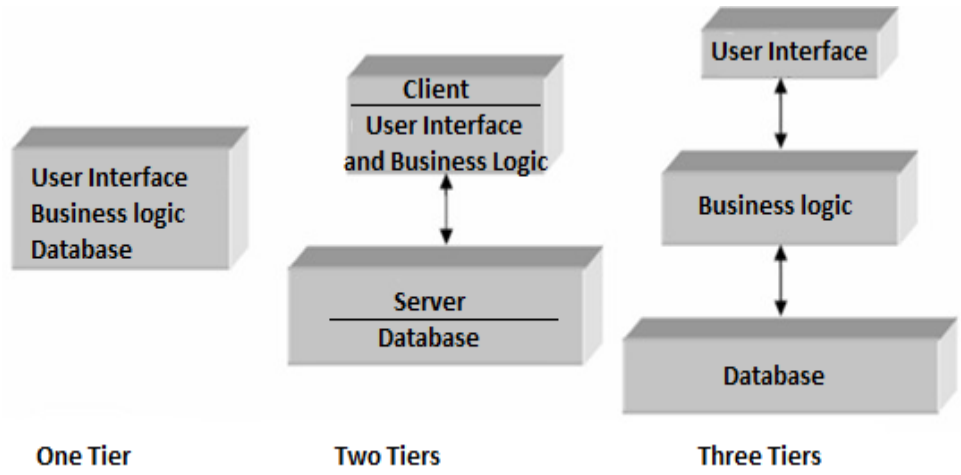


Figure 2. Tiers in the client-server information systems

The architecture of the classical web-based computing system includes four tiers (see fig. 3): Tier 1 includes the graphical user interface (GUI) visualization in browsers. Tier 2 includes GUI management realized on the web server. Within the scope of tier 3 is the business logic, implemented on an application server. Tier 4 includes the database management systems (DBMS) and all data sources used in the system.

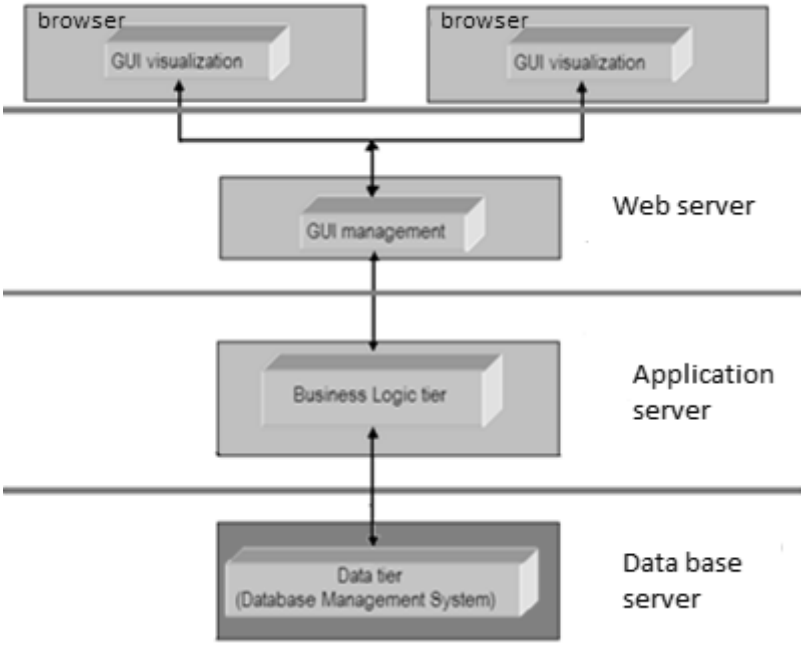


Figure 3. Four tiers in a classical web-based computing system

The major idea on which the transactional approach is based is that some set of actions performed by any end-user have to be performed in all-or-nothing manner (Gray, 1981). The aim of the transactional approach is to ensure consistency of the data in the database, given that many end-users work with one and the same set of data. While one user is reading some data, another user could try to modify or delete the same data. Such situations could be not resolved automatically by the database management systems. The transactional approach has an impact on the organization of data processing at all levels in the system and on the organization of the processes on the business logic tier, in particular.

Let us recall four principles of transactional approach in client/server systems, known as ACID. ACID stands for Atomicity, Concurrency, Independence and Durability:

- ***Atomicity:*** Each transaction is executed to completion, ***or not at all***? The term atomic is used to refer to operations that have multiple sub-operations but that are performed in all-or-nothing manner.

- ***Concurrency***: Transactions that are executed operate concurrently by sharing accessed data. Concurrency control has to ensure that transactions are performed without violating the data integrity of the database. The concurrency is needed for correctness in systems where two or more transactions, executed with time overlap, can access the same data. In cloud-based applications components are executed in concurrent processes. (Murdjeva, A., Stefanov, G., 2012).

- ***Independence***: Transactions are designed to operate independently of each other. Each client is ***written*** to operate, as if the entire remainder of the system were idle and the database server itself prevents concurrent transactions from observing one another's intermediate results. Such a process is referred to as isolation.

- ***Durability***: The results of committed transaction are persistent.

Nowadays many software developers consider the principles of ACID as absolutely necessary because these principles ensure the consistency of the data in the databases. The principle of data consistency in the database is regarded as fundamental. Furthermore, from the classical viewpoint, an information system with inconsistent database is seen as unusable at all. But there is some opportunities to avoid strong restrictions of ACID in classical client-server applications.

3. Conclusion

In the classical non-distributed client-server systems it is obvious to keep the principals of ACID. These systems are designed based on the ACID approach.

At its core, each non-distributed client-server system operates with atomic, concurrent, independent and durable transactions.

Distributed systems are different (Lazarova, V., 2012). There are another four principles that described processes more accurate but this are not considered in this article.

References:

1. Gray, J., 1981. The Transaction Concept: Virtues and Limitations. Proceedings of the 7th International Conference on Very Large Databases. Vallco Parkway, Cupertino CA
2. Lazarova, V., 2012. Multitenancy in Cloud Computing. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE, 5-6 October 2012, UNWE, Sofia, Bulgaria.
3. Murdjeva, A., Stefanov, G., 2012. Analysis of Conflicts during the Integration of Heterogeneous Multi-Dimensional Data Marts. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE, 5-6 October 2012, UNWE, Sofia, Bulgaria
4. Pritchett, D., 2008. BASE: An ACID Alternative. ACM QUEUE. [Online] Available at: http://queue.acm.org/detail.cfm?id=1394128 [Accessed December, 11th, 2012]
5. Radoev, M., 2012. Testing the Volume and Query Performance in Relational Databases in Case of Using Natural or Non Natural Keys. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE, 5-6 October 2012, UNWE, Sofia, Bulgaria.

Informal Economics of Information Threats

Serghei Ohrimenco[1], Grigori Borta[2]
[1]Laboratory of Information Security, Academy of Economic Studies of Moldova
osa@ase.md
[2]Laboratory of Information Security, Academy of Economic Studies of Moldova
grigori.borta@gmail.com

Abstract. This paper attempts to define shadow information economics as a domain of knowledge that aims at designing and implementing information threats (e.g. malware, DDoS attacks, etc.). This paper also analyzes and explores economical basis of shadow information economics functioning. An economical model of information threats is proposed.

Keywords: information security, information economics, shadow information economics.

1. Introduction

The phenomenon of shadow information economics is, according to our opinion, not sufficiently studied, even though it remains an important problem in the computer era, where cybercrime becomes a problem, that every user has come across. This paper tries to draw researchers' attention to the problem of shadow information economics.

2. Definition

We define shadow information economics as all the individual and collective unlawful activity, related to design, production, distribution, support, and use of components of information and communication technologies that is hidden from society. In other words, shadow information economics is all the criminal information products, services and processes based on IT or using IT. The main economical elements of this domain are unlawful economical relationships, illegal business, which is related to production, distribution and use of prohibited goods and services, sphere of illegal employment. It is important to note the fact that this kind of economics merges unlawful goods and services production, prohibited by national legislations, unlawful sale and purchase of goods and services, and consume of aforementioned unlawful goods and services. Therefore, we can conclude that the main reason of shadow economics existence is a set of conditions that makes it profitable to conduct unlawful activity in the domain of information technologies.

3. The Threats

A threat in information security is possible danger of a vulnerability being used to overcome system defense and cause damage. ISO 27005 defines a threat as follows: "a potential cause of an incident, that may result in harm of systems and organization". NIST defines a threat as: "Any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, or individuals through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service. Also, the potential for a threat-source to successfully exploit a particular information system vulnerability". Therefore, we can derive the following categories of threats:

- unauthorized access

- destruction
- disclosure
- modification of information
- denial of service

A research by Spy Ops, Technolytics, and Intelomics defines the following cyber threats:

- Logic Bomb
- Computer Virus
- Rabbit
- Bacterium
- Spoofing
- Sequential Scanning
- Dictionary Scanning
- Digital Snooping
- Spamming
- Tunneling
- Scavenging
- Counterfeit Equipment
- Counterfeit Software
- Software Malfunction
- Botnets
- Trap / Back Door
- TEDs / EPFCs / EMP
- Insider Threat
- Trojan Horse

The research defines a rating and a color code for each of the threats.

Threats may be classified by their type (physical damage, natural events, loss of essential services, information compromise, technical failures, and function compromise) and origin (deliberate, accidental, and environmental).

Another research, by Digital Forensics Association, covering 28 countries and 3700 incidents, shows that the main vectors of information breaches are hacks, removable storage, web, fraud (social engineering), and lost laptops.

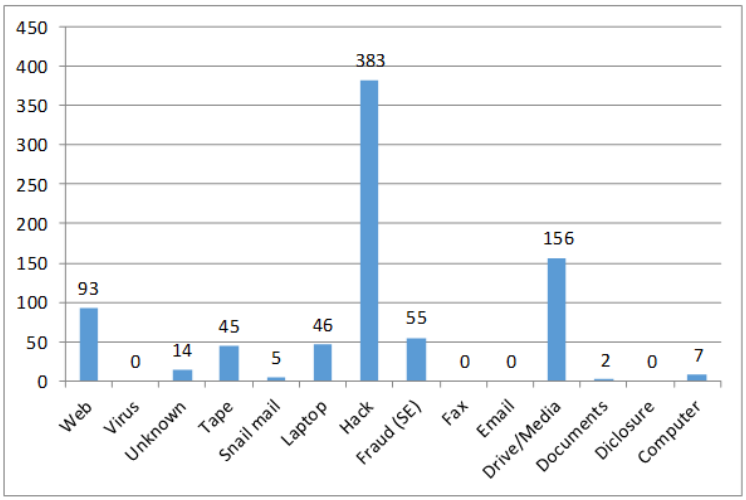


Figure 1. Information disclosure according to Digital Forensics Association.

A report by Symantec confirms that in 2012 the most usual cause of data breaches were hackers, being the cause in 40% cases, with accidental disclosure and loss of computers or drives being close second, both being the cause in 23% cases.

The same report suggests that the number of targeted attacks has grown by 42% compared to the previous year. This type of attacks more often targets smaller companies than before. According to Symantec, most of the attacks target manufacturing (24%), with Finance, Insurance & Real Estate and Services – Non-Traditional being on the second and third places, having 19 and 17 percent respectively.

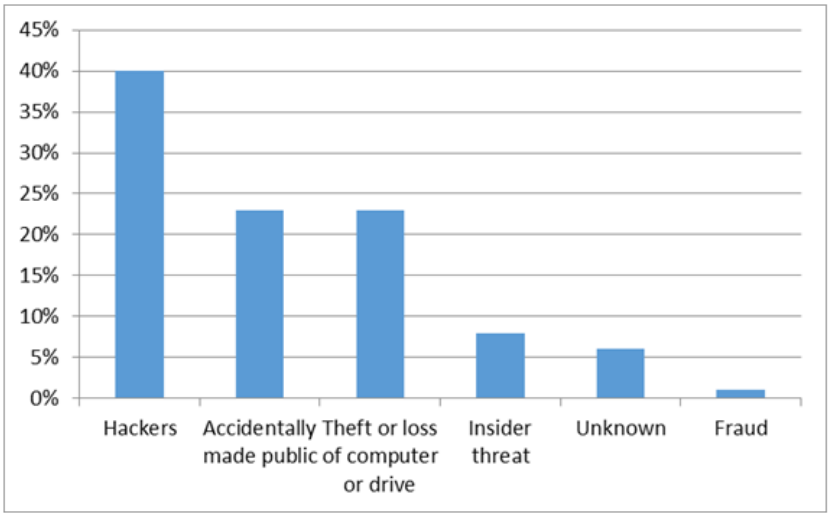


Figure 2. Data breach vectors in 2012 according to Symantec

4. Structure

In the Table 1 we provide a comparison between classical legal economics, classical shadow economics and shadow economics in the information technologies.

One of the major differences between the three is the employed goods and services. In the domain of information technologies, goods and services may often be confused due to their intangibility.

Goods in economics often refer to something intended to satisfy consumer needs. Goods are often considered to be tangible, while in the domain of information technologies only computer hardware is tangible. The most common criminal information technologies goods are the following ones:

- Specialized software – even though a lot of software in this category may be created with good intentions such as penetration testing for information security professionals in mind, these tools may be often used without proper authorization and for malicious purposes.

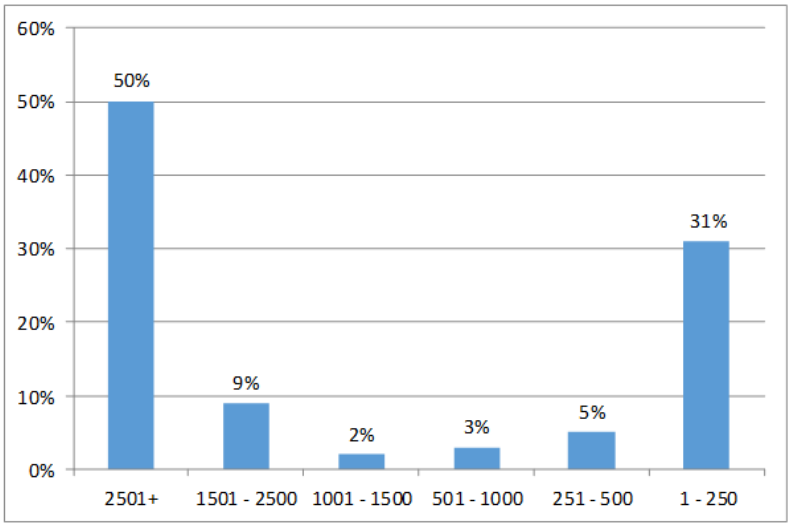


Figure 3. Number of attacks, depending on the enterprise employees number according to Symantec

- Spyware devices – credit card skimmers, electromagnetic impulse readers, hardware key-loggers, etc.
- Card counterfeiting equipment – fake plastic card producing and copying equipment.
- Pirated software – a major profit loss as is being claimed by software manufacturers.
- Private data – stolen logins, e-mails, passwords, credit card data, and all of the possible data, that may be used in identity fraud.
- Software and hardware vulnerabilities – according to Rainer Boehme and his paper named Vulnerability Markets, a stable and well-formed market of vulnerabilities exists. The researcher defines five types of vulnerability markets: bug challenges, bug auctions, vulnerability brokers, exploit derivatives, and cyber insurance.

Among the most common criminal information technologies services we would like to outline the following ones:

- Analytics – software vulnerabilities analysis, reverse-engineering, market and legislation analysis.
- Social Engineering – a method of acquiring personal data via winning one's confidence pretending to be a person that victim would normally trust.
- Spam – unsolicited advertisement delivery. Even though anti-spam techniques have substantially evolved during past years, undesirable messages still get through sometimes, which is annoying to users and may sometimes cause mailing systems malfunction.
- Pharming – attack via redirecting end user's web traffic from a legitimate web-site to a malicious one.

	Legal Economics	Shadow Economics	Shadow Information Economics
Goal	Satisfying consumer needs, maximizing profit using legal means	Satisfying destructive needs, maximizing illegal profit	Satisfying destructive needs, maximizing illegal profit
Relationship basis	Property relationship	Illegal use of property	Illegal use of property
Forms of property	• Private • State • Combined	• Private • State	• Private • State
Production, distribution, exchange and consumption	Related to forms of property	Illegal goods and services	Illegal goods and services
Methods and forms of activity	• Private • State • Combined	• Private • State	• Private • State
Management models	• Liberal • Neo-liberal • Keynesian • Japanese	• Arborescent	• Arborescent
Competition	Adjustable, intra-sector, inter-branch, combined (fair business conditions), imperfect competition (considerable restrictions apply), monopolistic (usually related to the uniqueness of goods)	Monopolistic, imperfect competition, self-regulated, overpriced goods	Monopolistic, imperfect competition, self-regulated, overpriced goods
Competition protection	State-level protection by law	None	None
Activity outcom	Legal goods and services	Illegal goods and services	Illegal informational goods and services

e			
Subjects	• Private persons • State	• Private persons • State	• Private persons • State
Objects	• Private persons • State	• Private persons • State	• Private persons • State
Types	• Profit • Non-profit	• Profit	• Profit
Interest	• Profit • Public service	• Profit	• Profit
Goods	• Private goods • Common goods • Club goods • Public goods	• Weaponry • Drugs • Human traffic • Illegally logged timber • Animals and animal products • Alcohol • Tobacco • Biological organs • Currency • Fuel • Counterfeit medicine, essential aircraft and automobile parts	• Specialized software • Spyware devices • Card counterfeiting equipment • Pirated software • Private data • Software and hardware vulnerabilities
Services	• Business functions • Childcare • Clear, repair and maintenance • Construction • Dispute resolution • Education • Entertainment • Financial services • Foodservice • Health care • Hospitality industry • Information services • Risk management • Social services • Transport • Public utility • Etc.	• Human traffic • Racketeering • Transportation providers	• Analytics • Social Engineering • Spam • Phasing • Pharming • Extortion • Sabotage • Terrorism • Piracy • Proxy services rent • DoS attacks • Botnet creation and rent • Money laundering
Employment	Official	Unofficial, unaccounted	Unofficial, unaccounted

- Phishing – attack aimed at acquiring personal data via tricking user into entering his personal data (such as login information, passwords, e-mails, credit card data) on a malicious web-site pretending to be legitimate one, which the user would normally trust.
- Extortion – being very similar to real-world extortion, its virtual-world counterpart aims at rendering some of the user's hardware useless unless a certain amount of money is paid. It is not unusual that even after the sum is transferred, malware continues to extort.
- Sabotage – deliberate action aiming at causing system malfunction.

- Terrorism – it is difficult to provide an unambiguous definition of this term, but, generally, a terrorist is a person who uses violence and coercion as means of reaching his or her goals. When speaking of terrorism in the domain of information technologies, one would often imply DoS attacks, web-site defacing, sabotage, specialized malware. Hacktivism may also fall into this category, even though it is usually much less harmful.
- Piracy – a service aiming at copyright infringement via unauthorized software copying, distribution, etc.
- Proxy services rent – a large proxy network may be often used to prevent malefactor's physical detection while sending patches and commands to botnets.
- DoS attacks – often sold as a service to those who don't have enough resources to perform it on their own.
- Botnet creation and rent – botnets may be widely used for DoS attacks, spamming, as proxies, etc.
- Money laundering – by means of information technologies. It is noteworthy that not only money earned in criminal sector of information technologies may be laundered this way. And after being laundered it may be used to further elaborate and develop both cyber and real-life crime.

5. Conclusion

The authors are aware of the complexity and complicacy of the issues brought up in this article, and realize that no quick solutions are available. Heightened interest to this subject gives hope for a successful resolution of the problems, but many issues are still left unresolved.

We find it important to draw researchers' attention to the following question, which requires immediate attention: What is the solution to the problem of shadow information economics? How is it possible to defeat it? Is it possible to completely annihilate it? Probably, not. Is it possible to lead all the shadow entities out the underground? One might suggest using the stick and carrot policy, punishing everyone who doesn't, and stimulate those who do. Or should we try and create the circumstances, where being in the shadow would be economically disadvantageous?

References

1. Spy Ops, Technolytics, Intelomics. *Cyber Weapons Threat Matrix*, http://spy-ops.com/web/CyberWeapons_Threat_Matrix.pdf
2. Symantec, *State of Information Global Results*, 2012, http://www.symantec.com/content/en/us/about/media/pdfs/2012-state-of-information-global.en-us.pdf
3. Digital Forensics Association, *The Leaking Vault 2011 Six Years of Data Breaches*, 2011, http://www.digitalforensicsassociation.org/storage/The_Leaking_Vault_2011-Six_Years_of_Data_Breaches.pdf
4. Andy Greenberg, Forbes. *Shopping For Zero-Days: A Price List For Hackers' Secret Software Exploits*, http://www.forbes.com/sites/andygreenberg/2012/03/23/shopping-for-zero-days-an-price-list-for-hackers-secret-software-exploits/
5. Boehme, Rainer. *Vulnerability Markets. What is the economic value of a zero-day exploit?* http://events.ccc.de/congress/2005/fahrplan/attachments/542-Boehme2005_22C3_VulnerabilityMarkets.pdf
6. Raj Samani, McAfee. *Cybercrime Exposed, Cybercrime-as-a-Service* http://www.mcafee.com/us/resources/white-papers/wp-cybercrime-exposed.pdf
7. Symantec Corporation. *Internet Security Threat Report 2013*, http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v18_2012_21291018.en-us.pdf

SWOT Analysis of Cloud Computing

EMIL DENCHEV
Department "Information Technologies and Communications" - UNWE

Abstract. The report was carried out as a SWOT analysis of the possibilities for cloud computing – Infrastructure as a Services (IaaS), Platform as a Services (PaaS), Software as a Services (SaaS), Data as a Services (DaaS), Communication as a Services (CaaS) and Monitoring as a Services (MaaS). Discussed are the strengths and weaknesses, and threats to this modern information technology based on the Internet. Emphasis is on the use of "good" practices and effective protection.

Keywords - SWOT analysis, IaaS, PaaS, SaaS, DaaS, CaaS, MaaS.

"Cloud" services (Cloud Computing) are installed on a web server maintained by the service provider and the access is done through a web browser or mobile applications.

"Cloud" services allow users to easily and securely access, store, process, synchronize and share computing resources across the Internet through a variety of devices such as PCs, smart phones, tablets, PDA (Personal Digital Assistant) devices, television and more.

According to experts from the consulting firm Protiviti "the advantages of the cloud model are obvious - companies save up to 60 % of their expenditure on IT infrastructure ".

SWOT analysis is a method to assess the Strengths and the Weaknesses, the Opportunities and the Threats available in the "cloud" services.

Under analysis are both internal and external factors in order to achieve the goals of the organization.

Strengths

Functionality

Cloud services allow you to store and process any kind of content, just as it is done on the hard drive - from text documents to music and video files. Some services allow you to store emails, contacts and e-calendar. Services allow accessing, editing and sharing content, no matter what platform, desktop or laptop, smart phone or tablet is used. Other important feature is synchronizing files from all devices. Sharing is protected by the usage of passwords and encrypted files.

Mobile Access

One of the biggest advantages of cloud services can be the access from different devices. Regardless of the device and the location - service (at work) or personal computer (at home), smart phone or tablet (on the fly), cloud services allow you to use certain content across a large number of applications and browsers for mobile devices.

Ease of use

Mobile applications and browsers that allow the use of „cloud" services are intuitive and has user-friendly interface. Mobile applications are easy to be configured and to be setup for the needs of each customer. What remains to be done are only local settings through predefined templates.

Help and Support

Get help with a problem when using the "cloud" services through communication with a telephone, via email or chat. The service provides a forum and other information materials as well as the services of highly qualified and experienced IT professionals.

Maintenance, including updating and upgrading the content of the services is performed by the supplier. This saves time and costs, as for example in business information systems, where the process of updating and upgrading is a complex and expensive process.

Reduced costs

The use of "cloud" services reduces the costs associated with the purchase of hardware, software, room hire and qualified staff to care for their maintenance. Companies pay only services ordered and used, i.e. customers pay a monthly fee and are not required to purchase expensive licenses. The time required for analysis and management of IT resources is diverted to key business activities.

Example: since no investments are required in ICT infrastructure disk storage system for archives (back up) with a high degree of reliability and performance can be provided. The prices and the complexity of backup and archiving models increase in proportion to the amount and the importance of the data.

Flexibility

The provider of a "cloud" service offers the ability to change at any time the necessary IT resources to respond to a change (growth or decline) in the business environment;

Weaknesses

User data is not on the local disk

Storing user data in a remote location (data server) is not a problem but an advantage as there are centralized access anywhere and easy archiving. However, if you use a service system for enterprise resource planning (ERP system) data is downloaded nightly in Business Intelligence (BI system) to be ready for the analysis of business` managers - obvious lack of easily and quickly accessible data is a serious problem. Still the speed of the Internet for the ordinary businesses is much lower than the speed of communication between the servers of the company and will take a long time.

Apply different security policies

Ordinary security policies on the local system are different from the security policies of the provider of "cloud" services. This means that the created users in the Local Active Directory are not automatically created in the "cloud" application, passwords are managed separately and there is no centralized monitoring of logs (log security).

Difference in functionality offered

"Cloud" applications may have different functionality (more or less) than the one offered by local applications. This can lead to the presence of "excess" functions or to certain limitations in the operation and utilization of the application.

Opportunities

"Cloud" services are divided into the following types:

1. Infrastructure as a Service (IaaS),
2. Platform as a Service (PaaS),
3. Software as a Service (SaaS),
4. Monitoring as a Service (MaaS);

Infrastructure as a Service (IaaS)

The IT infrastructure used is already in place in and the customer pays only for the resources that are necessary. Thus, the user can "create" server in a virtual environment without any problems and limitations that exist in the physical installation of hardware. This means that you can at any time ~~to~~ create, start, stop and remove a server. The infrastructure as a service includes:

- Data warehousing as a service (DWaaS) - the supplier provides information for use in direct marketing campaign and business intelligence;
- Data backup (Backup as a Service) - the provider creates and stores backups data and, if necessary, restores the status thereof;
- Communications as a Service (CaaS) – the supplier provides communication channels for fast and safe (protected) access.

Platform as a Service (PaaS)

Set of tools for software development and software applications that are located on the server of the provider and accessed via the Internet, regardless of the operating system, the resources, etc. Example of hiring a comprehensive platform for the development of software is Windows Azure, including C # + Windows Azure Compute + WCF + ASP.NET MVC + Azure Tables + SQL Azure + Azure Blobs + Azure CDN.

Software as a Service (SaaS)

Use of software resources through a portal interface (front-end), which practically guarantees their accessibility from anywhere on the planet. Such services can (include both) vary from web-based email to full use of the integrated management system of the business or just individual components. [2]

Examples of SaaS services:
• Enterprise resource planning systems (ERP),
• Customer relationships management systems (CRM),
• Supply chain management systems (SCM),
• Human resource management systems (HRM);

Monitoring as a Service (MaaS)

Provider provides services for monitoring and analysis of applied IT policies - for example, security policies, etc;

Threats

To consumers of "cloud" services there are certain threats mostly related to security.

• In many of the "cloud" services developers do not comply with encryption protocol that protects data from unauthorized access. Various mobile applications have different levels of protection - as well as the level of security and in response to new attacks.

• Vendors are developing new updates for mobile applications, and new mobile applications to reduce risks and make their services more secure, but this requires

downloading and installing these applications, which carries risks of viruses. In the future it is expected that "cloud" services will provide a unified solution that will increase protection against attacks.

Companies in Cloud Business

1	Microsoft Azure	10	Cisco Cloud Applications
2	IBM Cloud	11	Intel Hybrid Cloud
3	Apple iCloud	12	Dell Cloud Computing
4	Oracle Public Cloud	13	Adobe Creative Cloud
5	SAP NetWeaver on Demand	14	CA Cloud Solutions
6	Google App Engine	15	Symantec cloud services
7	Amazon Web Services	16	EMC Atmos Cloud Delivery Platform
8	HP Cloud Services	17	Salesforce Force.com Cloud Computing Platform
9	VMware Cloud Foundry	18	The Rackspace Cloud

Companies offering SaaS services in Bulgaria are:

1. Bulgaria Salesforce (salesforce.com) provides CRM system.
2. Bora Solutions (bora-bg.com) offers a "system of business management."
3. Intelligent Systems Bulgaria (isystems.bg) offers a CRM system from Microsoft.
4. Microinvest Bulgaria (microinvest.net) provides CRM software Sugar CRM.
5. Aloe Co. (erp.bg) provides ERP system Enterprise One.

References

1.Kristeva, N., Magazine CIO, ITC Media, Sofia, Issue 1, 2013
2.Denchev, E., Magazine CIO, ITC Media, Sofia, Issue 2, 2013

Cloud Information Technologies – Smart Business Analysis Solutions

Rositsa Ivanova
UNWE, 1700 Sofia, Bulgaria
Rosi_Ivanova@abv.bg

Abstract. Traditional information technologies (IT) for financial and business analysis of the organisation's business (business analysis) are based on the use of computer equipment and appropriate software. The fast information and communication technologies' development, clients' increasing demands, competition, globalization, world economic and financial crisis are the main factors that affect the necessity for development and introduction of new IT solutions in the business analysis field. There is an objective process of transition from traditional IT to cloud technologies for business analysis. The report highlights key issues related to the use of cloud technologies for business analysis, their weaknesses and strengths, as well as the new options they provide for the analysis of the organization's entire business in operational and strategic aspect.

Keywords: information technologies, business analysis, cloud technologies, operational analysis, strategic analysis

1. Introduction

The traditional model of information technologies in business is based on an overall computer infrastructure for information processing maintained by the company for the needs of its management. They can be defined as a system of methods, instructions, guidance and descriptions of procedures for the performance of the information processes within the organization: registration, storage, processing, transfer, delivery and use of information.

The financial and economic crisis that has occurred and is being developed has a direct impact on the use of computer technologies at global level. The crisis raised a new type of thinking for owners, managers and financial managers, who tend to shift the use of their own computer infrastructure with the use of external IT service provider, i.e. cloud service provider.

The cloud is a new model computer technology use. Cloud services could be defined as a set of hardware, software, network and know-how being used for storage, processing and use of information of remote computers in real time via internet. Therefore, the cloud may be considered as business software (SaaS), platforms (PaaS) and infrastructure (IaaS).

The starting point is that all organizations – internet users, can make use of the opportunities offered by the cloud services for processing and obtaining of information required to make correct and reasonable business decisions. Indeed, the shift in thinking is of special importance here, in addition to the understanding of the reality that information technologies, including cloud technologies, are a constituent element of the organization's overall business cycle.

The introduction of the cloud business model of using computer technologies will be a competitive advantage not only for the large, but also for the small and medium-size organizations.

This publication highlights the opportunities for using modern information technologies in economy. In particular, it outlines some major issues on the use of cloud technologies for business analysis, their strengths and weaknesses, as well as the new

opportunities they offer for the analysis, in operational and strategic aspects, of the organization's overall business.

2. Body text

The financial economic analysis (business analysis) is understood and treated within its interrelated and justified expression as a science, as a management function, and as a practical activity. Business analysis is an economic science with its own place in the scientific space, with its own subject, object, method, methodology, technique, functions, objectives, tasks and principles. It is an expression of the special management function of "analyzing" within the management system of every organization. As a practical activity, the business analysis is a subsystem of the organization's financial management system.

The business analysis is aimed at studying the activities (innovation, investing, main (according to the organization's scope of business), trade and financing), performed by the organization; the capitals (equity and borrowings) invested in these activities, and the specific property forms (non-current and current assets) they are embodied in; the financial results and the efficiency of these activities being reflected in the information system; the indicators for characterizing them and the factors justifying any deviations from set criteria level, for the purposes of objective assessment and making adequate and reasonable management decisions about the future activity, the organization's financial and market stability and its prosperity.

The analysis and assessment of the strategic and operational goals set by the organisation for the internal business processes and their efficiency, for the financial position and the financial results, for the efficiency of the cash flows, are based on the modern information technologies (see Figure 1).

Business analysis may be classified as internal and external. This differentiation is made also with view of the possibility to use cloud information technologies for business analysis.

The internal business analysis is characterized as follows:

- There is objective integration between forecasting, reporting, analyzing and making management decisions as special management function within the organisation's management system.
- The analysis is performed on wide information base, i.e. it may use public information, as well as internal information, which is not accessible to outside users, for example: different internal reports and forms, which are not obligatory for all organisations. We should note, however, that in practice there are some discrepancies when presenting the information created within the organization and required for making a profound analysis. For example, the formulas used to calculate the average amount of fixed tangible assets or short-term assets may be specified. By using different formulas, different values of those assets' average amounts are defined. Therefore, different indicators are defined, which, for example, characterize the efficiency of fixed tangible asset use and the short-term asset convertibility.
- The analysis' methodology and technique are complex and comprehensive and encompass a large number of analysis objects, by identifying and assessing the factors' impact on the behaviour and the dynamics of efficient indicators that characterize different aspects of the organisation's activity, both separately and in their mutual relations and dependences.

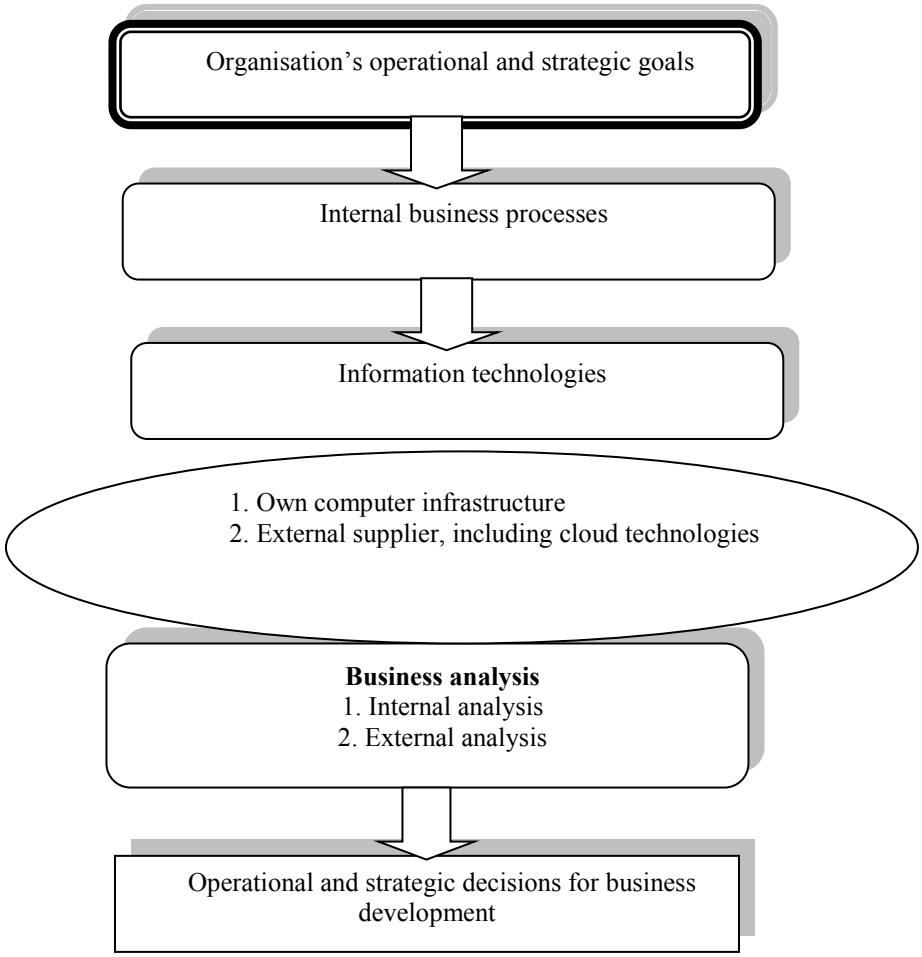


Figure 1. Business analysis

- The efficient analytical information is intended for the needs of the managing body (the organisation's management, owners, shareholders) to develop and make appropriate management decisions, i.e. the creation and use of this information is intended for intraorganisational use only.
- The efficient analytical information is characterized with a certain degree of confidentiality and is aimed at keeping the trade secret.

The specific features of the external business analysis may be systemized as follows:

- The analysis is performed on the basis of information from the organisation's financial statements, i.e. on the public external reporting.
- The elements of the organisation's financial statements are typical for all organizations that prepare and present their financial statements on the basis of the International Accounting Standards and the National Financial Reporting Standards for

Small and Medium Size Enterprises, thus providing conditions for, and allowing, the unification of the external analysis' methodology and technique.

- The aim of the financial statements with common purpose is to provide information about the financial position, the results from the activity and its efficiency, the efficiency of resource use, the equity and its dynamics, as well as the organisation's cash flows.
- The external analysis' methodology and technique are based on the assumption for information limitation.
- The efficient analytical information is public and is required for a number of internal and external (creditors, investors, clients, suppliers, state authorities, etc.) users with different objectives. This information is necessary for the users in order to make appropriate and reasonable management decisions.

On the basis of the above, we can make the conclusion that there are three important and objectively existing differences between the internal and the external analysis. The first one is in respect to the scope and availability of the information required for the analysis, the second one – to the analysis' methodology and technique, and the third one – to the contents, structure and availability of efficient analytical information provided to the different users of such information.

Understanding the differences between the internal and external analysis is the starting point when choosing information technologies for the organisation's activity business analysis.

The users of information (both internal and external) from the organisation's financial statements should obtain correct, fair and authentic information about the financial position, the financial results, the efficiency and the cash flows.

The organisation's financial position may be analysed in the following directions:

1. Analysis of the passive (liability) expression of value, including analysis of the content, structure and dynamics of capital – both equity and borrowings; analysis of capital – equity and borrowings, convertibility.
2. Analysis of the active (asset) expression of the value, including analysis of the content, structure and dynamics of the assets – fixed and short-term assets; analysis of short-term asset convertibility.
3. Analysis of the active-passive (asset-liability) balance sheet correlations, including solvency and liquidity analysis, capital reality analysis.
4. Analysis of the organisation's financial stability, including a system of absolute and relative analysis indicators.
5. Diagnostics of the financial position and defining the model of the organisation's financial position.
6. Analysis of the risk of trade insolvency and bankruptcy of the organization.

The modern approach for the organisation's financial position analysis is based on the objective integration between the management and the economic approach (see Figure 2).

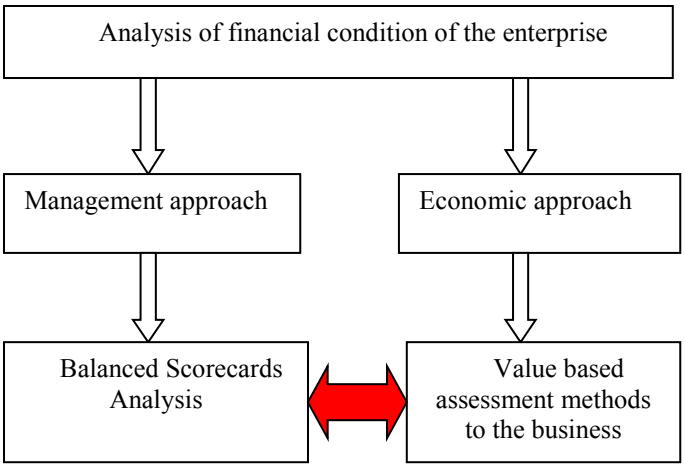


Figure 2. Financial position analysis approaches

The management approach is based on the integration of the balances system of financial (profit, profitability, etc.) and non-financial (management quality, ability to achieve the selected strategy, etc.) indicators for analysis of the organisation's efficiency and financial position.

The economic approach for the organisation's financial position analysis is based on the use of financial indicators for value assessment of business (value based business assessment models).

The organisation's financial results may be analysed depending on the way they are identified. Here we may distinguish separate methods for analysis of gross profit, respectively loss (accounting financial result), book profit, respectively loss, profits, respectively losses, from usual activity, from financial income and expense, and last but not least, the sales profit, respectively loss. The analysis of the accounting and the book financial result is an external analysis and is based on the data from the Profit and Loss Account, the Income Statement, respectively, (or the Statement of Comprehensive Income), while the analysis of the sales financial result is strictly internal analysis and no information about such analysis is disclosed in the organisation's financial statements.

The business efficiency analysis may be performed in different directions, such as analysis of the efficiency of use of resources held and controlled by the organisation. This analysis is an internal analysis and is based on the use of variable information sources, such as forecast, accounting, operational, technological, normative, and other types of information. The profitability, reviewed in its capacity of summarizing relative efficiency indication, may be analysed in different directions, depending on the bases chosen for its calculation, for example expense, cost, revenue, sales revenue, assets, production assets, capital, equity. To our view, it is substantial part the profitability analysis is an external analysis. However, profitability analysis based on cost, as well as on production capital (assets) gives important information to the internal users and may be rather considered an internal analysis. Furthermore, equity profitability and different analysis models give useful

and important information to owners about the funds they have invested, i.e. this is strictly internal analysis.

The organisation's cash flow analysis may be performed in different directions, such as availability, use and efficiency of their use. As far as the cash flow information in presented in the Statement of Cash Flows, such analysis may be defined as external analysis. To our opinion, the forecasting analysis of cash availability in the organisation in general and by types of activities, the use and efficiency of the cash use (profitability, liquidity, turnover) is rather internal analysis. The efficient analytical information is an important source of data for the managing subject. It can be used as a base to make appropriate business decisions for optimisation of cash content, structure and dynamics, for timely identification of any negative fluctuations of such cash, for development of effectively operating system of preventive measures to avoid liquidity crisis, both in short- and long-term aspect.

The analysis of the business plan performance, by scope, product range, quality, complexity, competitiveness and steadiness, is strictly internal analysis.

Furthermore, the expense analysis by types of activities, including by calculation items, as well as the cost analysis, are internal analyses as well, and these methods give important information for different types of expenses in different activities (main activity, sales, administrative activity) and for the cost on the basis of which different business decisions can be made, such as decisions for the development of new product, for eliminating a product, which is not sufficiently effective, from the product list, etc.

The general integrated business analysis approach is based on the objective symbiosis between the management and economic approaches (see Figure 3).

Based on modern business analysis approaches, methodologies and techniques, we definitely believe that the time has come for organisations' officers and managers to look for perspective proposals ensured by modern information technologies. And why not the innovation idea to use cloud technology services for business analysis?

Share of resources, realization of material economies, as well as establishment of extremely favourable business environment are some of the strengths of cloud calculations, which make them fully applicable in economy, and in particular in business in relation to the organisation's activity analysis. The specific application of cloud technologies for business analysis depends on the set goals and tasks, as well as on the type of analysis – internal or external analysis. Currently, the business in Bulgaria shows reserved attitude to the investments in cloud technologies due to many different reasons. Here are some of the main reasons: fear and mistrust to information security in multi-subscription environment; mistrust to the safety of technologies and information against unauthorised access; fear of viruses, which might block the systems and the information; fear and mistrust to the reliability of equipment and communications; apprehension about data exchange speed in real time; lack of skilful staff both for technical support and for development of new applications.

Based on the above, we believe that the cloud services for business analysis of the organisation's activity may be used by establishment of internal and external cloud.

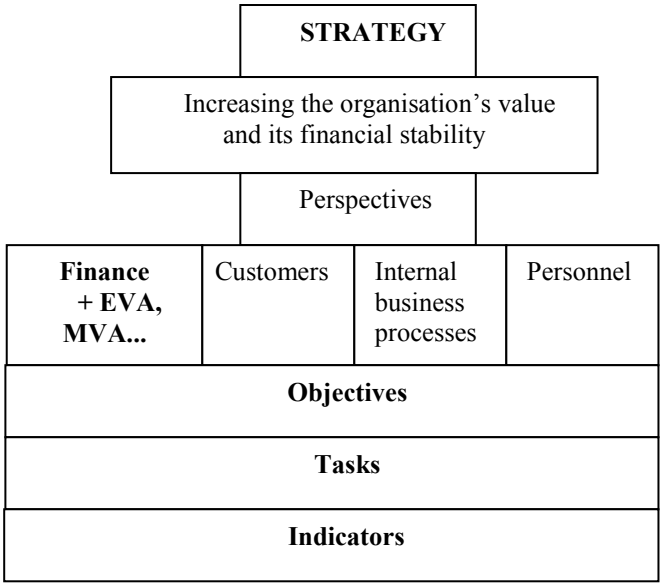


Figure 3. Integrated business analysis approaches

The internal cloud is more effective way to organise the data processing centre, but it requires completely new IT architecture. The development of a cloud calculation model for realisation of internal cloud should start with evaluation of the IT equipment homogeneity, which is used within the organisation, by virtualisation of servers through the virtual machine on the basis of developed images. Within the IT cloud specialists adjust the process for construction of the virtual machine once, and then all processes are fully automated. The access to the internal cloud is ensured within the frames of the organisation's corporate network, both at one point (office, units, etc. according to the functions of the organisation's responsibility centres), as well as in remote points (branches, affiliates, suppliers, distributors and other business partners). Each organisation has the opportunity to objectively evaluate all facts and to authorise or deny access to the organisation's internal cloud for any information users from remote points.

The internal cloud is an open system and modern information technology users may take advantage of the possibilities external providers of different ready applications offer. Thus, the opportunities for outsourcing of information-technological structure of the organisation in general, and of the business analysis, as a special function of its management, significantly increase.

Most organisations consider the idea for establishment of internal cloud more reasonable, since they believe the external cloud cannot ensure the respective high level of information security that might be achieved with the information systems with traditional structure. Thus some issues rise for the managers involved in the business information provision. Currently, they are not ready to give up the control of the organisation's

information-technological infrastructure, while at the same time the "cloud service" market is still in process of establishment.

When using generally accessible external clouds for business analysis we should take in consideration both the organisation's goals, tasks and requirements, as well as the forecasts about their use. Here it is very important to know in details the services offered by different cloud service providers, to make critical comparative analysis and objective evaluation of their weaknesses and strengths, whereas the choice of supplier should be bound by contract stipulating the respective provisions that fully protect the interests of the organisation-user of external cloud services.

The organisation may get advantage of a mixed option of internal cloud with access to external cloud. Internal clouds may be developed, which develop applications to be exported in an external cloud. We believe that such cloud technological structure may be applied when making a business analysis of the organisation's activity.

Our opinion is that the internal cloud can be successfully applied for an internal business analysis, and the external one – for an external business analysis.

The next step in the development of cloud services and their use in economy, including for business analysis of the organisation's activity, is to develop a private cloud of different types:

- Private cloud developed by the organisation's data centre, including software, hardware and services;
- Private cloud developed by the organisation's data centre located within the organisation but controlled by external provider;
- Private cloud developed by the provider's data centre located at the provider's and controlled by him.

The strengths of the use of cloud services for business analysis can be systemised as follows:

1. Optimisation of organisation's expenses. The expenses incurred for establishment and maintenance of IT infrastructure (computers, servers, disk arrays, communication equipment, air conditioning, power supply, etc.) have a significant relative share in the aggregate expenses for the organisation's activity. Here the benefit and expense convention objectively applies. According to this convention, the benefits from the use of efficient analytical information should exceed the expenses incurred by the organisation for the development and the provision of this information to the managing subject. By using cloud calculation procedures for business analysis, organisation's expenses are significantly reduced, while the benefits for the users of business information exceed the expenses for its development, transmission and use.

2. Efficiency optimisation. The efficiency of using the traditional IT infrastructure in the organisation is low due to the impact of a number of factors, such as power cut-off, which causes losses to the organisation in addition to all lost profits, since errors may occur in the information systems or servers, as well as opportunities for hacker attacks and thefts of business information about the organisation's development in operational and strategic aspect.

3. Optimisation of the risk level. The application of analytical methodology and technique in the calculation cloud (the business analysis is considered a service within the cloud) optimises the risk levels in the following ways:

- The service is a low risk model to apply specialised resources through subscription;

- The service is transparent, since it is standardised and characterised with extremely high level of reliability and quality;
- The organisation is not interested in the way this service is provided, as well as in the resources it is provided with and at what price;
- The service is a highly automated process which reduces the impact of any subjective factors, such as human errors, etc., to the minimum extent;
- The technological risk level is reduced to the minimum extent, and in most cases is totally eliminated;
- Continuous business is ensured as a result of the continuity of the technologies.

4. Optimisation of information security. Information security is one of the most discussed issues when developing the organisation's information infrastructure. We should note here that security is not universal, but is rather a set of physical, logical and information structure. Information security in information technologies is achieved when confidentiality, integrity and availability of data are ensured. With the traditional approach to information technology use, small and medium-size organisations in Bulgaria forecast in their budgets only expenses for achieving physical information security. In order to guarantee the availability of data at any time and to ensure continuity of business processes, they are required to provide backup equipment in case of failure of any of the computer system's elements. When using cloud technologies for business analysis, a high level of information security is achieved in its different aspects.

5. Optimisation of business development perspectives. The use of cloud calculation procedures for business analysis gives rise to various perspectives for the organisation originating from the inexhaustible technical resources they have. For example, the internal business processes may be optimised, comprehensive analyses of financial position, financial result, efficiency and cash flows may be performed, appropriate management decisions for the business development in operational and strategic aspects may be made. On this basis, the organisation may achieve financial and market stability, as well as improved competitiveness and prosperity.

6. The use of cloud services is a prerequisite for the development of mobile and flexible business. Internal resources are released within the organisation, objective opportunities are established to generate new business ideas, to offer new products and services on the market, to increase the efficiency of staff use by means of improving its motivation and satisfaction, thus creating competitive advantages.

3. Conclusion

Currently, the managers and owners of organisations are oriented to growth strategies. They focus their attention to increase their competitive power, to keep their loyal clients and to attract new clients, to conquer new markets, to open and cover new market niches, and thus to increase the organisation's market share. In this way they can maintain the organisations' financial and market stability and prosperity of their business.

Organisations should focus on using modern information technologies, including cloud technologies, in their business, as means of control, quality assurance, production capacity improvement and keeping loyal clients.

References

1. Technologic, "Information technology for manufacturing companies", 2012
2. Electronic document, http://idg.bg/stone/news/41.

CLOUD COMPUTING VERIFICATION OF HYPOTHESIS VIA A COMPLEX: STATISTIC ESTIMATIONS – DISCRETE DISTRIBUTIONS

Lubomir H. Lahtchev, PhD

Institute of Systems Engineering and Robotics, Bulgarian Academy of Sciences,
e-mail: lahchev@iser.bas.bg

Abstract: The mathematic statistics contains a number of methods about tests and verification the hypothesis. The Z-statistic is disseminated as a popular method of test and verification the hypothesis at enough large aggregate arrays. Characterized peculiarities of these methods are the integral estimations and their mutual dependence, which allows checking the final decision. At the same time, the real distribution of a random value by data of representative aggregate arrays stands away the attention. Actually the aggregate arrays can be processed and discrete distributions can be calculated. Several cloud computing estimations can be incorporated on this aim. The comparison of discrete distributions on overlapping of their graphics and analysis can serve for verification of a real hypothesis represented in the paper.

Key words: verification, hypothesis, discrete distribution, statistic, test.

INTRODUCTION

The mathematic expectation, dispersion and mean squared deviation are the often met statistic estimations at test of an aggregate array. The comparison of random values with issues of pairs of data is an actual scientific task solved with use of cover and correlation functions, method of the coefficient of mutual correlation. The verification of a hypothesis about several random values is accomplished by method of Z-statistics [3, 9] at enough large aggregate arrays. Actually these methods fulfill the clouds of the information cloud computing approach and are a base for introduction of some new trends in the statistics. In parallel, different cloud computing techniques applied on one and the same aggregate arrays generate different solutions. Our experience showed that statistic estimations issued from the statistic distribution by repeatability are larger the same taken for data of equal probability. As a result cloud of assumptions imposes requirements, which fluent the final values of the estimations. The difference is quantitative only, when one and the same aggregate arrays are assessed. In order to improve the quality of statistic processing the discrete distribution is suggested as a detail, visual, valuable graphic estimation of a random value, represented by aggregate arrays. The recommendations of yesterday about empirical distribution of random value would be reassessed. For example, Korin in [9], chapter 5 writes: "Before specifying the sampling distribution for the sample mean we shall construct a frequency distribution for sample mean. While generating this empirical distribution is again not necessary, doing so should clarify the concept of sampling distribution".

Sage A.P. and Melse J.L. [4] write about some difficulty of discrete distributions: "… usually the mean value or cover function of a random process are quite easier found than the law of distribution". But the law of distribution represented by corresponding function on discrete values of aggregate arrays is affordable cloud computing algorithm at test and verification some hypothesis with random values.

The aim of the current paper is test and verification of hypothesis:

$H_0 : \mu_{UA} = \mu_{FA}$ "The BP in the upper arm and forearm are equal"

and its alternatives:

$H_1 : \mu_{UA} <> \mu_{FA}$ "The BP in the upper arm and forearm are not equal"

with use of the main statistic estimations and discrete distributions by deviations of second order.

The selected aim can be achieved with a package cloud computing processing of the data about aggregate arrays with the known statistic variables and multi-threshold separation of the ranges of data by statistic deviation of second order inside the area of distribution.

CLOUD OF THE ASSUMPTIONS

The measurements of BP with one and the same sphygmomanometer are completed in right and left upper arms by a scheme, shown in [2], in right and left forearms, and with record of data about systolic and diastolic BP. The cuff over upper arm is installed in accordance to the recommendations of the producer of the sphygmomanometer. The cuff over the forearms would be installed in a way preventing its sliding – the air hose is going toward the upper arm. As a result 8 aggregate arrays with N=81 data in every one of them or 648 data are available. Each one pair of systolic and diastolic issues about BP is independent result of measurement. Hence, the preliminary computation of the statistic distribution by repeatability of the data issues in accordance to the equations:

$$E\{(x=\alpha_i)\} = (1/N)* \sum_{i=1}^{N} (\alpha_i * n_x(\alpha_i)) = \sum_{i=1}^{M} (\alpha_i * p_x(\alpha_i)) \qquad (1)$$

where n_x is number of issues at $x = \alpha_i$ and $p_x = n_x/N$, i – number of groups with equal issues, M – number of all issues after the processing on repeatability, falls out. The equations 1 clarify the data frequency processing of statistic variables of the random values and some explanation in the corresponding formula in [1].

The data about BP are unique or of equal probability – 1/N = 1/81.

Pursued condition of stationary of random processes, described by the aggregate arrays, is accepted in general by default. It is scientifically proven about the upper arms in another issue [8].

For simplicity the outliers are neglected.

The each one array is processed and average or mathematic expectation, deviation of first order, dispersion, and deviation of second order are produced.

Discrete distributions are assessed by mathematic expectation and mean squared statistic deviation.

The verification of the hypothesis of two random values represents comparing the graphics of the corresponding discrete distributions and analysis by overlapping.

The role of the Z-statistics at verification of the hypothesis is also explored.

CLOUD OF THE STATISTIC ESTIMATIONS

With an account of the equal probability, the mathematic expectation represents known average. It can vary from one aggregate array to another as result of available outliers and data clusters.

$$E\{(x=\alpha_i)\} = \mu_x = (1/N)* \sum_{i=1}^{N} \alpha_i \qquad (2)$$

In the current exploration the mean squared deviation is as follows:

$$\bar{} \quad N \quad \bar{} \quad N$$

$$\sigma = (1/N) * \{\sum_{i=1} sqrt[(\alpha_i - X)^2 \ . \ p_x(\alpha)]\} = (1/N) * \sum_{i=1} \sigma_i \qquad (3)$$

The function of discrete distribution in the practice of the mathematic statistics is often substituted by Z-statistics, based on the following equation:

$$Z = (\mu_x - \mu_y) \ / \ \{sqrt[D(X - Y)]\} \qquad (4)$$

The parameter Z is computed for the random values at estimation, which in the current explorations are BP in the upper arms and BP in the forearms. With an account of the level of significance noted by α for an aggregate array the function of distribution Φ_o is defined by the following equation [3]:

$$\Phi_o(Z_{\alpha/2}) = (1 - \alpha) \ / \ 2 \qquad (5)$$

The standard table of the function $\Phi_o(Z_{\alpha/2})$ [3] allows definition of the corresponding values $\pm Z_{\alpha/2}$, that represent the boundaries of critique ranges.

The H_o hypothesis is supported by low values of $Z < Z_{\alpha/2}$. From one point of view, the alternatives of H_o or hypotheses H_1 can be expressed by comparisons of the mathematic expectations of the compared random values. A special interest at the current exploration is verification of the results about a hypothesis, which is described in [7] by general and differential probabilities – an innovation alternative approach to testing a hypothesis.

DISCRETE DISTRIBUTIONS AT VERIFICATION THE HYPOTHESIS

The function of discrete distribution $f(x_{ij})$ represents the i-number of issues corresponding to a j-range of the dispersion of the random value. It can be computed for first, second and higher orders of the deviation the random value. In the current exploration the second order of deviation is chosen.

Discrete thresholds define the ranges of distributions of the random values and they include also supported by the author algebraic signs: $\pm L_1 = \pm 1.\sigma$, $\pm L_2 = \pm 2.\sigma$, $\pm L_3 = \pm 3.\sigma$, $L_4 = 4.\sigma$.

With an account of differences $(\sigma_i - \bar{\sigma})$ the function of discrete distribution of an aggregate array on the discrete thresholds is built on the following logic expression:

$$\text{If } (\sigma_i - \bar{\sigma}) \leq L_j \text{ then } f(\alpha_{ij}) = f(\alpha_{ij}) + 1 \qquad (6)$$

where α_{ij} – j-th discrete range with i-issues inside it. As a result of processing it is established independence of the discrete function of distribution on the chronology of the data or some other data ordering.

Discrete distributions of different random values are compared by essential details and summarize parameters: shape coefficient and shape distance to a centre. The parameters can be used at different algorithms for assessments in accordance to chosen principles.

The shape coefficient is an empirical characteristic of the discrete function calculated by the following equation:

$$C_f = [\sum_{k=0}^{I-1} \sum_{j=-Lj}^{+Lj} (i_0 + k*\Delta i) * f_s(x_{ij})] \ / \ [S(i,j)] \qquad (7)$$

where i_0 = const, Δi =const, S(i,j) – surface of the form given by discrete distribution.

The shape distance to a centre is a length of ort connecting the central moment line of the discrete distribution with the ordinate axis. It can be computed by analytic or graphic processing of the function of discrete distribution, similarly to the formulae about C_f.

EXPERIMENTAL RESULTS OF VERIFICATION BY Z-STATISTICS

The hypothesis that "The blood pressure within the forearms exceeds the blood pressure within the upper arms" is estimated with a probability of 72.2222% in [7]. In the current exploration it is verified by Z-statistics. The table 1 shows results of verification the hypothesis H_o "The blood pressure in the upper arm and forearm are equal" by quintiles of Z and $Z_{\alpha/2}$.

Table 1 Verification of the hypothesis H_o by Z-quintiles.

Parameter of BP at	Level of significance α	Function $\Phi_o(Z_{\alpha/2})$	Quintile $\pm Z_{\alpha/2}$ of Φ_o	Quintile of array Z	Inference
Right arm Systolic BP	0.05	0.475	1.96	3.592	Reject H_0
Right arm diastolic BP	0.065	0.4678	1.85	5.414	Reject H_0
Left arm Systolic BP	0.02	0.48899	2.29	2.114	Accept H_0
Left arm diastolic BP	0.055	0.4798	2.5	4.3	Reject H_0

It is seen 75% of rejection the hypothesis H_o. Consequently, the author's test in [7] is verified positively and the values of the mathematic expectations on table 3 show higher BP in the forearms in comparison to the BP in the upper arms.

EXPERIMENTAL VERIFICATION OF HYPOTHESIS VIA A COMPLEX: STATISTIC ESTIMATIONS – DISCRETE DISTRIBUTIONS

The theory advises that the real distribution of random value is given by the law of distribution the probabilities, corresponding to the issues. With the introduction the signed discrete thresholds discrete distribution acquires some other quantitative definition. Thus each detail of the discrete distribution can be assessed and compared to the same one for another random value or aggregate array. Moreover, it can be estimated by summarize parameters of shape coefficient and shape distance to a centre. Table 2 contains important information about the all graphic discrete distributions of the explored random values about BP.

Table 2. Discrete distributions of random values by aggregate arrays.

Location and signal	-2	-1	0	1	2	3	4
Upper arm right, Sys	0.0	0.00000	0.59259	0.32099	0.07407	0.00000	0.0
Upper arm right, Dias	.0	0. 00000	0. 55556	0. 34568	0. 09877	0. 00000	.0
Upper arm left, Sys	.0	0.00000	0.54321	0.30864	0.13580	0.00000	0.0
Upper arm left, Dias	.0	0.00000	0.60494	0.27160	0.11111	0.01235	0.0
Forearm right, Sys	.0	0.20988	0.39506	0.18519	0.11111	0.09877	0.0
Forearm right, Dias	.0	0.07407	0.49383	0.30864	0.09877	0.02469	0.0
Forearm left, Sys	.0	0.22222	0.37037	0.25926	0.09877	0.00000	0.0
Forearm left, Dias	.0	0.03704	0.51852	0.33333	0.07407	0.02469	0.0

The careful sight on the table 2 recognizes narrow distribution of the random values about BP within the upper arm and increased dispersion of them within the forearms. The corresponding graphics follow on the next fig. 1 about the right arm by systolic and diastolic BP. The same ones for the left arm are shown on the fig. 2. Their forms are close to the normal distribution and demonstrate a good quality of the graphics about the random values.

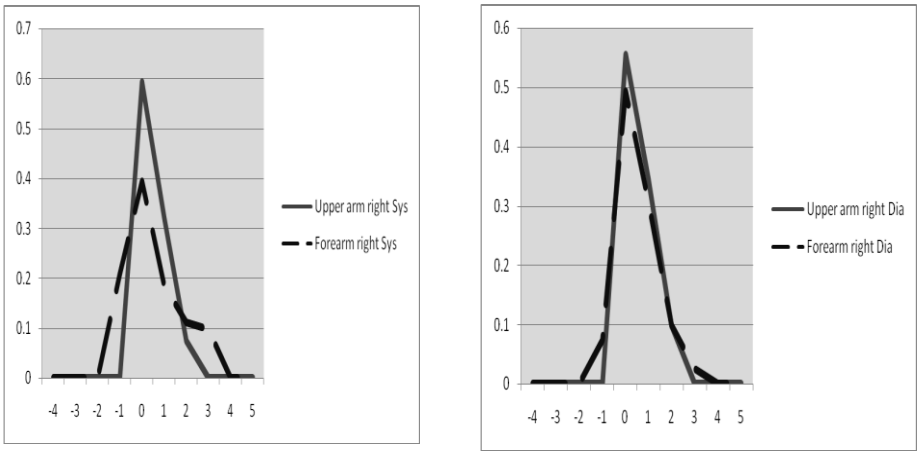


Figure1. Discrete distributions of the right arm by systolic and diastolic BP.

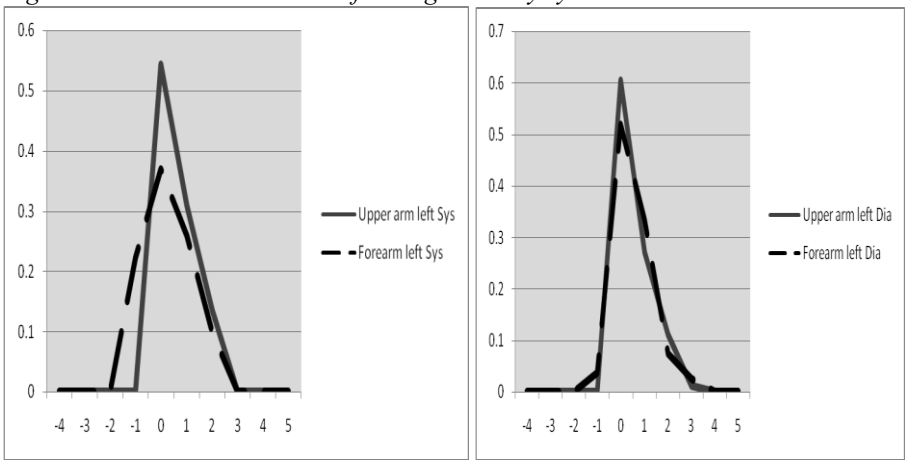


Figure 2. Discrete distributions of the left arm by systolic and diastolic BP.

The important differences of the graphics about systolic BP at both right and left arms are features about higher dispersion of the corresponding random values generated for forearms. The similarity of the same features and the corresponding differences at the small ranges about the diastolic BP are witness for specific distributions of the functions of them.

The analytic assessments of the graphics are completed by the shape coefficient, shape distance to a centre and surfaces of the forms under the lines of distributions. They are shown on the table 3.

Table 3. Features of estimations the discrete distributions.

Location	Mean value	Disper-sion	Mean squared deviation	Shape coeffi-cient	Shape distance to centre	Surface
Upper arm right, Sys	122.22222	73.78875	0.75781	0.27110	0.36840	0.73860
Upper arm right, Dias	79.20988	54.07026	0.65872	0.28850	0.47500	0.64250
Upper arm left, Sys	119.48148	64.66758	0.68536	0.23650	0.55000	0.66790
Upper arm left, Dias	76.95062	50.00381	0.61927	0.23230	0.53100	0.61900
Forearm right, Sys	127.95062	132.29767	0.98127	0.17820	0.45000	0.97020
Forearm right, Dias	86.07407	75.98445	0.75751	0.18834	0.47000	0.74800
Forearm left, Sys	122.80247	129.60844	0.94212	0.30000	0.1750	0.91845
Forearm left, Dias	82.16049	68.66819	0.72685	0.21242	-0.4686	0.71830

The higher shape coefficient corresponds to higher level of BP. The preferable shape distance would be less than mean squared deviation. If the feature "surface" is accepted as estimation, then the BP in forearm totally prevails over the BP in upper arm. But it will not be true because higher surface at large boundaries of the discrete distribution supposes higher dispersion of a random value.

With comparison of the shape coefficients it is seen the prevailing at three locations of measurement of the upper arm over the forearm, and one location of prevailing for the systolic BP about left forearm over the left upper arm. This circumstance substantiates an analysis extra. When BP within upper arm is accepted higher than BP within forearm, then the mean growth deviation is 12.24%. In parallel, mean relative similarity is 87.32%. Hence, as a result of comparison the quantitative assessments of above suppositions, the hypothesis H_o can be accepted as true. This inference is supported by the increased dispersions of the corresponding discrete distributions about systolic BP at the forearms, assessed by deviations of second order. Consequently, differences of the BP at both upper arms and forearms come from the increased dynamics of the BP and of its corresponding signals at the forearms. With account of the discrete distributions of the BP, built by deviations of first order, the picture

can look different. Some qualitative features about verification of hypotheses by deviation of first order are already described in [8].

With account of the upper analysis the following recommendations can be suggested.

The test of hypothesis, which depends on the static parameter – mean value of the aggregate array of a random value, is recommended to be verified with Z-statistics.

The test of hypothesis, which depends on the speed of the estimated parameter – deviation of first order of the aggregate array of a random value, is recommended to be verified on graphics of discrete distribution, based on deviation of first order and on valuable estimations of it.

The test of hypothesis, which depends on the dynamics of second and higher order of the estimated parameter – deviation of second and higher order of the aggregate array of a random value, is recommended to be verified on graphics of discrete distribution at large boundaries, based on deviation of second and higher order and on valuable estimations of it.

CONCLUSION

The verification of hypothesis via a complex: statistic estimations – discrete distributions reveals a number of variety at the corresponding information technology by cloud computing algorithms. It is shown that this complex provides abundance of useful information about the explored random values. The current exploration infers a specific separation of the methods about test and verification of hypotheses. It is result of comparisons about phenomena, observed at cloud computing the data arrays about BP. The results are computed with use of Excel.

REFERENCES

1. L. Lahtchev Cloud Computing Individual Assessment of a Medical Device for the Education. Conf. Proc. of the Intern. Conf. on Appl. of Inf. and Comm. Tech. in Economy and Education, December – 2011, UNWE, Sofia, Bulgaria, pp. 503 – 512.

2. L. Lahtchev. Cloud Computing Statistics of Individual Alternative Tests with a Sphygmomanometer. Conf. Proc. of the Intern Conf. on Application of Inf. and Comm. Tech. and Statistics in Economy and Education. October 2012, UNWE, Sofia, Bulgaria, pp. 322-330.

3. И. Бронштейн, К. Семендяев. Справочник по математике для инженеров и учащихся втузов. Лейпциг, „Тойбнер", 1979, Москва, „Наука", 1980 г., 975 с.

4. Э. Сейдж, Дж. Мелс. Теория оценивания и ее применение в связи и управлении. Москва, „Связь", 1976, 496 с.

5. E. O'Brien; Th. Pickering; R. Asmar; M. Myers; G. Parati; J. Staessen; Th. Mengden; Y. Imai; B. Waeber; P. Palatini; W. Gerin; N. Atkins. *International Protocol for Validation of Blood Pressure Measuring Devices in Adults.* Blood Pressure Monitor., February 2002, vol.7, issue 1, pp. 3-17.

6. P. Hartmann. *Validation of the Hartmann Tensoval Duo Control (DS1881) device according to the British Hypertension Society protocol in an adult population.* Published on the site: http://es.hartmann.info/ES/Inicio/Los_blank_Productos/Incontinencia/validacion_de_BHS.pdf , visited 20.11.2009

7. L.Lahtchev Likely-hood-based Educative Data Analysis of Alternative Tests with a Sphygmomanometer. Conf. Proc. of the Intern Conf. on Application of Inf. and Comm. Tech. and Statistics in Economy and Education. October 2012, UNWE, Sofia, Bulgaria, pp.

8. L.Lahtchev Verification of Hypotheses about Blood Pressure by Statistic Estimations and Distributions. Proc. of Intern. Conf. On Automatics and Informatics'2013, October 3- 7, 2013. Editor E. Nikolov. Publ. By John Atanassoff Society of Automatics and Informatics, Bulgaria, 2013, pp. I-33 – I-36.

9. Korin, B.P., Introduction to Statistical Methods, Wintrop Publishers, Inc., Cambridge, 1977, 426 p.

Combining Virtualization Technologies in SOA-applications

Mihail Radev[1], Yanka Alexandrova[2]

[1] University of Economics - Varna, 77 Kniaz Boris I Str., 9000 Varna, Bulgaria
radev@ue-varna.bg

[2] University of Economics - Varna, 77 Kniaz Boris I Str., 9000 Varna, Bulgaria
yalexandrova@ue-varna.bg

Abstract. The paper examines possible aspects of applying virtualization technologies during the processes of development, testing, implementation and management of SOA-applications so as to achieve high quality software, shorter release cycles and easier modification and adaptation. The virtualization aspects examined are service virtualization, virtual services and abstract layer of virtual services. The authors point at the main advantages of each of the examined technologies during development and maintenance processes connected to SOA-applications. The comparative analysis of these three technologies concludes with recommendations about the collaborative application of all three aspects of virtualization in order to combine the benefits of each virtualization technology.

Keywords. Virtualization, Service Oriented Architecture, virtual services, SOA-applications.

1. Introduction

Current SOA application developers are facing many challenges and have to deal with a great number of problems, such as complicated project management, high pressure for short release cycle, increased demand for better software quality, different kinds of constraints and limitations.

The purpose of this report is to analyze and compare different virtualization technologies in order to discover their benefits and limitations regarding SOA applications development and eventually to propose recommendations for their collaborative implementation.

2. Virtualization methods and technologies

A major part of problems developers experience are due to different kind of limitations and obstacles, such as unavailable or with limited access operational systems, hard to simulate complex systems, scheduling conflicts, constantly changing requirements, and so on.

A recent research, made from Coleman Parkes and Computer Associates amongst 501 professionals in software development area in North America and Europe, reveals some of the major challenges, software developers have to face: increased demand for new software release (average 6.4 per year); higher requirements for more and better functionality; shorter release cycles (from 3 to 5 months per version) and strict budget limitations. At the same time respondents point out the main obstacles in software development and testing process. These include - undeveloped applications (marked as a significant challenge by 61% from

American and 39% from European software developers), limited or no access to third-party systems and databases, corporative mainframes and additional license costs associated with using outside resources only for testing and development purposes.

The cited research results show negative tendencies for those companies which use conventional approach in development, integration and testing. Although it is not stated specifically, these results are also relevant to SOA applications development as this area is characterized with elaborate production environment, complicated team collaboration and complex government.

We think that a significant part from the mentioned limitations and challenges could be overcome through using virtualization in different aspects. These aspects include:

1. Server virtualization (classic or hardware virtualization). From SOA point of view this virtualization aspect could be viewed as a service producer virtualization.
2. Service virtualization – modelling service communication and behavior.
3. Abstract layer of virtual services allowing one service implementation to be represented by several virtual services.

Server virtualization

Server virtualization helps to achieve consolidation of physical servers, lower maintenance costs and TCO. From software producers' point of view there are additional and more important advantages following server virtualization such as opportunities for application testing in different operational environments, server resource clustering, providing availability and reliability of services, government and control of services' load and system restore. Due to the nature of the current SOA applications and inconsistent resource load, one of the main advantages of server virtualization is connected with the possibility of dynamic resource allocation.

Service Virtualization

Service Virtualization is a practice of "capturing and simulating the behavior, data and performance characteristics of dependent systems and deploying a Virtual Service that represents the dependent system without any constraints, thus allowing software to be developed and delivered faster with lower costs and higher reliability". We have to say that the term "service virtualization" does not refer only to virtualization of web services, but also includes virtualization of all resources used in composite application development as service, data, mainframes, service bus, etc.

Server virtualization gives a good number of advantages but is applicable only to under allocated systems. As mentioned before the main part of constraints and problems are due to over allocated resources, which are hard to simulate through hardware virtualization or are unavailable in development in testing phases.

The main differences between server and service virtualization could be summarized in the following comparative table (see table 1).

Table 1. Comparison of server and service virtualization.

Server virtualization	Service virtualization
Increased usage of under allocated systems.	Decreased usage of over allocated systems.
Virtual image of accessible systems.	Access to systems, which cannot be virtualized by server virtualization.
Goal: virtualization of the whole system.	Goal: Simulation of system's behavior and communication.
Uses virtual machines	Uses virtual services

The comparison reveals that service virtualization could be used in cases where server virtualization is not applicable, e.g. when dealing with over allocated or inaccessible resources. This leads to the conclusion that the two types of virtualization are not interchangeable, but complementary.

Virtualization of a service is a process of creating a virtual service which consists of instructions for simulation of a communication between systems. The process typically involves three main phases:

1. Capture. A software component, called "listener" records messages and data exchanged between the application under development and the system which is being simulated.

2. Model. A service virtualization software correlates the captured data into a virtual service that may simulate a credible conversation of requests and responses.

3. Simulate. The deployed virtual service (VS) is used as a stand-in for the simulated system. VS will respond in a more predictable manner than the "real" system giving the appropriate data according to user requests.

The idea of simulating external systems is not new and has been widely adopted in process of development of complex, distributed and heterogenic applications. The traditional approach for simulation includes creating a "mock" software application, programmed to return certain data as a response to requests. The main disadvantage of this approach, however, is the narrow scope of possible answers generated from the prototype. This is caused by the limited scenarios of conversation situations coded in software prototype. In order to imitate broader set of communication scenarios developers have to put in significant efforts commeasurable with efforts needed to develop a new system. Another drawback of this conventional simulation approach is the impossibility to use it at the development and testing processes which may lead eventually to a later discovery or ignoring serious mistakes or omissions.

The service virtualization approach could help avoid the aforementioned problems through providing plausible environment in development and testing, automation of the process of simulation of external systems, facilitation of parallel team work and dynamic virtual service government through processing of exceptions and adapting VS model.

Abstract layer of virtual services.

The term "virtual service" used in this virtualization method has rather different meaning compared to the previous virtualization technique. Virtual service here plays the role of intermediate between the real implementation of a service and service consumers. Consumers interact with a virtual service with alike, but not exactly the same interface as that of the "real" service. The virtual service's interface may define attributes, descriptions and requirements, determined by different service contract in respect of service consumers' needs and requirements. The interaction between service consumers and producers is carried out in abstract layer which is responsible for synchronization of virtual services interfaces,

transformation if needed, integration with other infrastructures and policy-based changes. Virtual services form the abstract layer that has different addresses from the real service implementation, which allows modifications in service producer to be made without affecting service consumers (see fig.1).

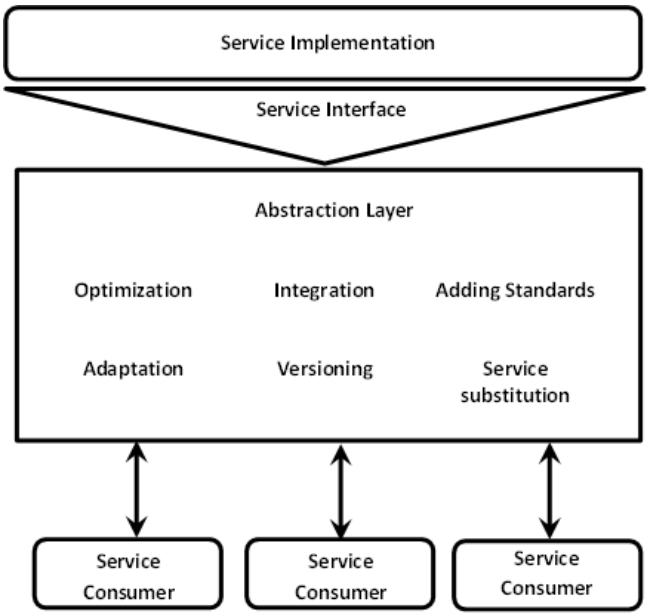


Fig.1. Abstract layer with virtual services

Virtual services in abstract layer give several advantages in the process of development of SOA applications such as:

- **Service versioning.** The loose coupling, achieved in virtual services, allows changes to be made in the abstract layer without impact on the service consumers. This permits different versions of one service to exist simultaneously. Representation of different service versions as virtual services is a more efficient approach compared to implementing service versions as different service implementations help minimizing the government and development efforts and increase service reuse which is one of the basic SOA principles.
- **Service replacement and termination.** The migration to new service producer or termination of services could be eased significantly if the needed transformations are made in abstract layer rather than in service consumers.
- **Introducing new standards or their versions.** There is a constantly increasing number of standards in web service domain. Every new standard, subset or standard version requires certain transformations to be made both to producers' and consumers' side. With virtual services the introduction and implementation of new standards could be carried out in the abstract layer thus avoiding changes and transformations in service consumers or producers;
- **Migration to new technology.** Composite applications are dynamic and constantly evolve as a response to changes in external and internal for the

company factors as legal regulations, requirements from partners (customers, suppliers, distributors) and others. These changes demand new technologies to be implemented and this requires a complete service makeover. This could be avoided if new technology implementation is made in the abstract layer;

- **Adapting to requirements of different service consumers.** In order to achieve good service reusability developers must meet the requirements of broad set of users who probably use different identification or security methods. The abstract layer could manage these requirements and perform the needed transformations and monitoring of services without changes in their implementation.
- **Integration of different SOA technologies.** Web services can be deployed using different technologies and protocols as Java Message Service (JMS), IBM WebSphere MQ, Representational State Transfer (REST, REST full), SOAP, SMTP (Simple Mail Transfer Protocol), and so on. The collaboration between these technologies and protocols requires significant efforts, which could be realized in the abstract layer thus contributing to increased service reusability and eased development and government.

The abstract layer of virtual services permits to achieve agile and secure SOA applications and a higher degree of service reusability. The virtualization of services, however, is not limited only to providing an interface for interacting with the real service. To fulfill at a maximum level the opportunities of this virtualization concept the abstract layer has to carry out integration and synchronization functions as well as the needed service policies and transformations.

3. Comparison and collaborative implementation of Virtualization methods.

The aspects of virtualization presented give a clear advantage in building composite applications. It should be noted that despite the specifics, these three aspects of virtualization are not mutually exclusive and may advantageously be used in combination, taking into account the possibilities and limitations thereof, the main ones are shown in Table 2.

Table 2. Comparison of the three virtualization methods.

	Server virtualization	Service virtualization	Abstract layer of virtual services
Virtualization of:	Resources	Service behavior	Interaction between service producer and consumer
Limitations	Could not be implemented on over allocated or unavailable resources	Implementation only in development and testing phases	-
Advantages	Increased resources' usability coefficient; High economic effectiveness;	Short time for development and testing; Reduced costs; Better software quality.	Agility and security of applications; Increased service usability; Loose coupling between producers and consumers.

4. Conclusion

The comparison of benefits and limitations of virtualization aspects, presented in table 2 gives us the ground to propose the following recommendations regarding collaborative use of virtualization methods:

- First, use of virtual services at the stages of development and testing of SOA-applications. The implementation of such technologies can be gradual starting from strategically important for the company applications. The most significant limitation of this aspect of virtualization is its application only in the development, modification and testing, but as these are key steps in the process of building applications and given the dynamics of SOA- applications , we believe that the virtualization services are essential for quality and the functionality of the composite applications;
- •Second, use of virtual services and abstraction layer for all applications, regardless of their scope , number of implemented services and users . Building abstraction layer should be laid at the stage of defining the requirements for the new application and is a top priority for developers. The use of virtual services contributes to a high degree of realization of the two basic principles of SOA - Reusable Services and independence between consumers and providers of services that ultimately allow the construction of adaptive and functional applications.

Collaborative use of the three aspects of virtualization allows multiplication of the positive effects of server virtualization, virtualization services and abstraction layer of services. Server virtualization is a priority in the construction of all kinds of applications and companies have very rich experience in its use. In SOA- applications this method of virtualization could be also successfully applied to all resources, suppliers or consumers of services. However, we believe that the combination of the other two aspects of virtualization would have the greatest effect precisely on creating SOA- applications since both approaches apply methods of abstraction and reuse of services. Combining virtualization services and abstraction layer can be implemented as a virtualization of behavior of services from the abstract layer. Virtualization services could be deployed on the basis of captured communication (exchanges of messages) between the virtual services from the abstract services rather than the real service implementation. This combined approach will allow for multiple versions of one real service to be simulated using a broader set of communication scenarios.

Examined aspects of virtualization provide a number of advantages for building SOA-applications such as optimizing the use of resources, shortening the deadlines for development, quality improvement, easier and flexible service management, increase the degree of utilization of services and independence of consumers and service providers and others. We believe that the three methods could be used together in the process of building and operating of SOA- applications thus combining most of the opportunities offered by virtualization.

References

1. Ariola, W., C.Dunlop, Sevice Virtualization Business Drivers, Parasoft 2013, available at http://www.parasoft.com/service-virtualization, [accessed 1.12.2013]
2. CISCO Systems, Service Virtualization: Managing Change in a Service-Oriented Architecture. available at: http://www.cisco.com/en/US/prod/collateral/ contnetw/ps5719/ps7314/prod_white_ paper0900aecd806693c2 .html [accessed on 31.13.2013]
3. Coleman Parkes, 2012. North America Study: The Business Benefits of Service Virtualization - Free Report, available at http://servicevirtualization.com/study [accessed 20.11.2013]
4. Michelsen, J. & English, J., Service Virtualization: Reality Is Overrated, Apress 2012

THE WAYS OF COMBINING THE DESIGN PATTERNS

Marya Armyanova
Department "Informatica" of the UE-Varna

Abstract: In this article are reviewed the Gang of Four (GoF) patterns and possibilities for combining them in new patterns. For the purpose are presented and classified relationships between classical patterns. So design patterns to be combined and integrated in software development. Described are different classifications, types of patterns and their place in CASE. In particular, we consider Enterprise Architect (EA)'s way of creating, integrating patterns as necessary to correct design, and using UML to describe them.

Key words: Design patterns, Enterprise Architect, CASE, UML, Gang of Four

Introduction

Patterns, and in particular design patterns are promising software technologies which can accelerate the process of development and to improve the quality of the developed system. In the development of complex systems, software developers having to combine a great number of patterns, in order to cover the various aspects of the system - the storage of the data by the use of various algorithms for processing and calculation.

When developing applications design patterns are applied together. The goal is to create a flexible architecture subject to reuse. Furthermore interaction between patterns is achieved, which is much more than a sum of several patterns.

There are many types of patterns with different assignment, functions and levels of typing. Therefore, there are different pattern classifications based on different criteria. Classical design patterns, described by the "big four" - Eric Gamma, Richard Helm, Ralph Johnson and John Vlissides [11] are twenty-three in number. According to Gamma a design pattern is a "description of communicating objects and classes that are customized to solve a general design problem in a particular context" [11]. Also Patterns of Software Architecture are classified [10] and other patterns are united together in a classification of programming languages, which are suitable, such as C++. They are called Pattern Languages of Program Design [4, 5, 6, 16]. By classic patterns preferably are solved problems of a low level of designing of software. They offer software solutions that especially provide a possibility for code reuse and ease maintenance. One of their main goals is to ensure the development of software that is easy to maintain and modify. Thus, in its tracking is not required to be made large and fundamental changes. Therefore, in the classic pattern described by the group of four, the software adaptation to the new requirements should be implemented without significant reforms in existing software. So they carry flexible software that easily adapts to the changed requirements. In contrast, by the CASE Enterprise Architect (EA), the patterns provided are on a higher level of typing. They cover all levels of development and are modeling patterns which are on higher level .

Martin Fowler [10] and a number of other prominent software developers offering patterns for designing the next level, which do not focus on creating classes and interfaces. Their patterns are formed on the basis of the use of the components in a broad sense. They stand at the basis of EA architectural patterns. Refer to the commonly used solutions in different information systems. These patterns are still developing and differ from classical patterns in objectives and reasons for its occurrence.

Architectural patterns are based on the integration of many design decisions needed for the development of a system. These patterns apply Service Oriented Architecture (SOA),

REST , etc. [1]. They determine the manner of establishing the overall system by integrating the smaller components within the scope of the software application. When using SOA, business logic generalize to the service layer and the user interface interacts with the different program modules providing services. Through these modules provide implementation of different business processes involved in the software system. In REST API patterns using REST protocol to provide resources through the Internet.

Classifications of patterns

Combining patterns in a system is performed by relating them in different ways, which should be described in the documentation of the pattern. However, collections of patterns contain only a description of the relationship between the elements of the collection. In the different collections same relationship is realized and described differently. Just this contradiction determined difficulties in using the patterns as the description of the relationship is significant part of the definition of the patterns.

The classification of the relationship between the patterns is considered to facilitate the use of patterns. It presents the basic interactions between the patterns - a pattern uses another pattern, a pattern refines another pattern, or a pattern conflicts with another pattern. Also are covered some minor relationship, such as: pattern being similar to another pattern, or one pattern combining with another pattern.

Different collections of patterns represent the relationships between patterns differently. For example, in the described by Zimmer [20] patterns are considered three types of relationships - a pattern used another, a pattern pattern used another pattern or two patterns are similar. According to Design Patterns Smalltalk Companion [1] besides the three relationships, there's another type - a pattern specialisation to another. Patterns of Software Architecture [4] also propose three relationships types, two of them are different - a pattern can use, be used by, generalize, specialize, or provide an alternative to another pattern. Other collections of patterns present the union and reflexive relationship between patterns [16, 17].

In general, the main relationships between patterns can be grouped as follows:

Primary Relationships: Uses; Refines; Conflicts.

Secondary Relationships: Used by; Refined by; Variant; Variant Uses; Similar; Combines; Requires; Tiling; Sequence of; Elaboration.

The three primary relationships between patterns are presented in the classical design patterns - one pattern uses another pattern; a more specific pattern refines a more general pattern, and one pattern conflicts with another pattern when they both propose solutions to a similar problem.

There are also described a number of secondary relationships between patterns, such as a pattern being a variant of another pattern, two patterns being similar, or two patterns combining to solve a problem. They are described how they can be expressed in terms of the primary relationships.

The first main relationship is **uses**. It is very common. It is found in all collections of patterns. One pattern uses another pattern. In this relationship a pattern creates models which can be used from other patterns. This allows other patterns to perform their activities. Typically, the global pattern with a larger application in the system is using the smaller, local pattern. This relationship is used to create more complex patterns as a composition of simpler patterns. Riehle [17] considered the creation of a pattern, as any design pattern which is best described as the composition of further patterns. It presents a Model-View-Controller

pattern, which is obtained as a composition of the Observer, Strategy, and Composite patterns.

The pattern Composite combine objects to create new functionality. It configures the objects in a tree structure to represent hierarchies of the "parts of the whole." This enables users to treat in the same way as individual objects and configurations of objects.

The Observer pattern connects a View to its Model, the Strategy pattern arranges for a Controller object to handle input on behalf of a View, and the Composite pattern provides a hierarchy of Views. A View does not handle user input but leaves this to the Controller – thus the Controller is a Strategy for handling user input. Views can have Subviews which represent smaller parts of the user interface and which can have Subviews themselves – thus the View is a Component in the Composite pattern and different Views can be either Leafs or Composites. Fig. 1. illustrates this pattern relationship.

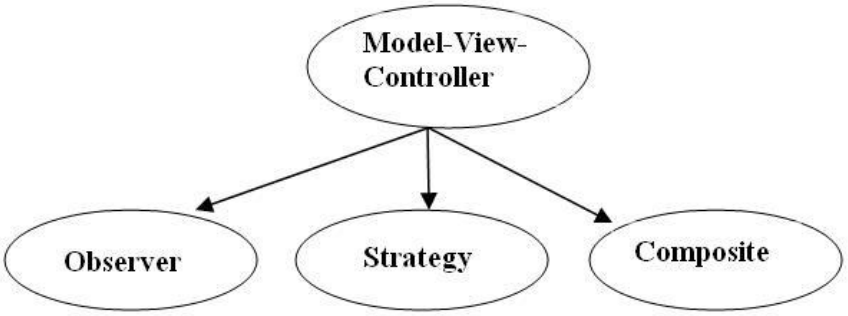


Fig. 1. Uses pattern relationship.

Other authors – Dyson and Anderson [16] are declared a collection of small patterns. Composite pattern uses the smaller-scale patterns of which it is composed. The description of a composite pattern must describe how these patterns are combined into the composite pattern.

Patterns often use this type of relationship in combination with others. In order to maintain consistency between different views of objects, the Observer pattern enters in this type of relationship with the Mediator pattern. So it can coordinate the update of objects. Turn the Mediator pattern also uses the Singleton pattern, to avoid unnecessary duplication of their sites.

The other major type of relationship is **refine**. In terms of object-oriented programming this relationship corresponds to the succession. The specific pattern is a setting of the general, incomplete or abstract pattern to a specific problem. The specific pattern adapted generic pattern to the environment. It offers a similar but closely specialized solution similar structure and a similar forces, but addresses additional forces.

This type of relationship is common for classic patterns, too. The Factory Method pattern refines the Hook Method pattern. By Hook Method pattern is allowing subclasses to override the behavior of superclasses of the appropriate pattern. Thus it is possible to comply with naming conventions, which facilitates future development of systems built. Factory Method lets subclasses to specify the class of an object in the superclass. Factory Method refines Template Method, because Factory Methods are effectively Hook Methods which are used by subclasses to specify the class of an object the Template Method in the superclass will create. This type of relationship is shown in Fig. 2.

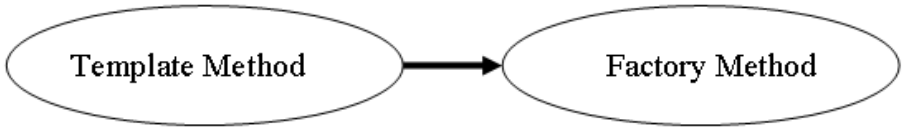


Fig. 2. Refine pattern relationship.

Through this relationship can define new patterns, as summarized existing ones. Thus created an abstract pattern, has a common structure and presents the general properties and behavior.

Another use of this relationship is at graphic systems. The Classic Iterator pattern refine the Curried Object abstract pattern. The Curried Object pattern defines the way in which the subject of the client application provides a simple interface to the object server application. For this purpose, the cumulative information in the local state about which object the server is using from the client application. This information is included to the message sent to the server. In fact, the Iterator pattern provides an interface for a set of objects by managing current position. Position of the object in the collection is added to the message. The similar patterns are presented by Design Patterns Smalltalk Companion [1] such as Sharing and OO Recursion, and it describes other patterns as refinements of these abstract patterns. For example, Flyweight refines Sharing, and Chain of Responsibility refines OO Recursion.

The **conflict** is the third fundamental way to connect the patterns. A pattern addresses the same problem as another pattern. It describes the case where several patterns offer mutually exclusive solutions to similar problems. In most cases, this relationship is not clearly described, but it is defined in the description of the related patterns, with the uses relationship, too. For example, two conflicting patterns which both address the problem of how a Smalltalk programmer should access an object's instance variables [15]. According to the Direct Variable Access the instances variables should be accessed directly, because this is easy to read and compiles efficiently. In contrast, the Indirect Variable Access pattern presents a diametrically opposed solution. The instances variables should only be accessed indirectly, via accessor methods, because this makes it easy to change the representation of the variable without affecting code that uses it. This relationship is presented in Fig. 3 .

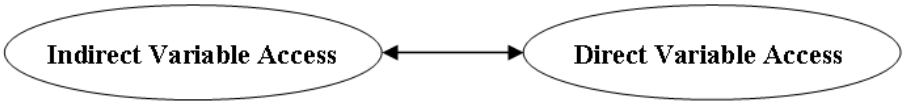


Fig. 3. Conflict pattern relationship.

In practice this type of relationship is applied in two ways. When the suitable pattern for the design is selecting, together with initially selected pattern should be investigated and alternative conflicting patterns. But in the end should be chosen and realized only one pattern.

In the collection of the classic pattern, the Decorator pattern conflicts with the Strategy pattern. Both are behavioral patterns (modify the behavior of other objects), but add graphical borders or icons to window objects in different window system frameworks. The Prototype and Factory Method patterns also conflict. They offer two different ways of how subclasses to redefine the classes of objects constituting the superclasses.

Unless used by the big four ways to connect patterns, a number of other authors offer and other relationships between patterns. These are considered as Secondary Relationships and are described by the primary relationships.

The first relationship of Secondary Relationships is **used by**. It means that a smaller pattern is used by a larger pattern. It is the inverse of the uses relationship and may be regarded as an alternative to it, for example, in the classic pattern Mediator uses Singleton. But the same relationship can be presented as a Singleton is used by Mediator. Or Iterator is used by Interpreter and Visitor. This point is important because different patterns in the implementation of the used by relationship save it in different sections, while others do not even recorded it.

Another classic feedback is **refined by** or generalises relationship. Therefore, it can be seen in the same context. A general pattern is refined by a specific pattern. For example, Template Method is refined by Factory Method.

Another type of the secondary relationship is a **variant**. A variant pattern refines a more well-known pattern. Since design pattern provide repeatable decision to common problems, they are abstract. When they are used, they must adapt to the specifics. Some specific conditions occur more often than others, so some ways of instantiating patterns are more common. This determines the creation of new patterns as variations of existing patterns. Some identify pattern variants, either as separate patterns or as parts of other main patterns.

There are different ways to create a variant of a pattern, which can be interpreted as new patterns. But not all are successful, widely used or important for practice. Some are substantially different from the main pattern, that is they extend the main pattern's structure or make different tradeoffs between forces, and, then they are worth documenting as variant patterns in their own right.

There are two distinct ways to use the variant relationship. Most pattern variants provide alternative solutions to the same problem. They are called solution variants. A few pattern variants describe ways a single pattern can provide a solution to a number of different problems. They are problem variants.

The next relationship is **similarity**. A pattern is similar to another pattern - "X is similar to Y" [17, 20]. This relationship includes relationships between patterns which are not captured by the other relationships. The patterns can be similar because they address the same problem, such as Abstract Factory, Prototype, and Builder. The similarity relationship is more general than the pattern's conflict. It also presents patterns which have a similar solution technique, such as Strategy and State, or which are occasionally used together, such as Flyweight and Strategy. The conditions of the particular implementation of the patterns are essential for the patterns identification as similar, unlike the primary relationships.

Another relationship is **combines**. Two patterns combine to solve a single problem which is not addressed directly by any pattern. An example of such a relationship has to happen when a pattern is wider than the decision of this problem, while the other pattern is narrower and solves only the part of the problem. Such is the case with the classic pattern Proxy and Forwarder-Receiver pattern. As the Forwarder-Receiver pattern modifies the Proxy pattern and implements communications substrate. In other words, the Proxy uses Forwarder-Receiver pattern.

This relationship is used when both the pattern can not solve this problem alone and they need the addition of new functionality. Provided that there is no pattern to solve the problem in its entirety, patterns should be combined. It must be ensured that these patterns cover the entire problem. Another option is to create a new pattern that is designed to solve

specific problems and uses the necessary decision patterns. For example, The Abstract Mechanistic Behaviour and Abstract Significant Roles patterns combine to implement an abstract design.

The next relationship is **requires**. A pattern requires the solution of another pattern if the second pattern is a precondition for solving the problem addressed by the first pattern. For example, the Software Framework, Abstract Significant Roles and Abstract Mechanistic Behaviour patterns all require the Defined Domain pattern - that is, the Defined Domain pattern must have been applied before the other patterns can be applied successfully [7]. The difference between this relationship and primary relationship uses is only in consistency of the patterns implementation. The required pattern must be applied first. If one pattern uses a second pattern, the second pattern must be applied before the solution described by the first pattern will be complete.

Another relationship is **tiling** [15], when a pattern uses itself. Sometimes some patterns are applied repeatedly until solve a particular problem. Such relationship typical examples are when Interpreter and Visitor patterns can be instantiated several times to implement a iterative system. Iterator can iterate over other iterators. Tiling patterns relationship is implemented by creating additional patterns templates, which are applied iteratively. In this context, the iterative pattern templates refines and uses the main pattern. A pattern which can be tiled is simply one which uses itself.

Using this relation shows that complex patterns can be tiled in different ways. For example the Composite pattern can be tiled in four different ways and the Cascade patterns are presented as a tree of composites where each layer in the tree contains different types of objects. Each of these variants both refines and uses the main Composite pattern.

The last reviewed type of relationship is **Sequences of Elaboration**. A sequence of patterns from the simple to the complex [10]. It is implemented as a sequence of patterns, which begins with small, simple, low-level patterns, which are united in complexity, architectural patterns. This type of relationship includes patterns from a fragment of a pattern language.

Sequences of elaboration can be represented by the primary relationships between the patterns in the sequence. For example, the larger patterns refine and use the smaller patterns in the sequence. Therefore, the more general and more applicable small scale patterns should be presented first. So sequences of elaboration should be presented beginning with the simple patterns, and then proceeding on to the more complex patterns.

EA's way of creating, integrating patterns

Patterns, and in particular design patterns are a promising software technology which can accelerate the process of development and improve the quality of the developed system. By CASE environments their uses facilitate and expand the scope of their application. EA combines many signages of the latest UML version with intuitively understandable interface. The method UML diagrams allow describe both the business model and the model of the entire software application, its internal organization and principles of the interaction of its parts [8]. Because of maintenance of different types of UML diagrams, it is possible a detailed description of the behavior of virtually any software application. Generation of code by developed UML diagrams is possible only for classes represented by class diagrams. Thus EA mainly focuses on the design of models of software applications rather than developing their program code. EA is using class diagrams and objects of the method UML to illustrate the structure of the pattern. Knowing the types of relationships between patterns allow designers to present them using UML diagrams.

The use of EA in the development of the system involves the development of a number of aspects: business architecture, information architecture, technology architecture, as well as the actual change management. Only a small part of the EA collection of patterns is focused on the design application. EA patterns are modeling, and some of them are design patterns.

EA patterns are another set of patterns that presents common solutions to frequent business problems. They cover certain patterns reflecting different aspects and levels of decomposition of software systems. For example, some of the architectural EA patterns used patterns Dependency Injection and Inversion of Control, to allow free combining different system components.

According to Martin Fowler [10] the pattern Dependency Injection is a special case of the pattern Inversion of Control. Dependency Injection, allows designers to remove the hard coded dependencies, i.e. inserted input or configuration data in the program code and allows changing them at run-time and compile. The pattern automatically inserts dependent element (object, value, etc.) on the allocated place by understanding and satisfying the requirements of the assignment itself. It is very suitable for dynamic testing.

EA patterns [3] are using Object Relational Mapper (ORM), to address the entity between relational databases and business objects. They differ in the components used in the construction of the different layers of business information systems.

The important, by the use of patterns in EA viewpoint, for development of business applications, is that they can be interpreted as a group of interrelated objects and classes obtained by the generalization of some of the possible business scenarios. For modeling business processes through UML [9] designers use activity diagrams, objects diagrams and personal profiles. EA adds UML 2.4.1 with BPMN (Business Process Model and Notation), which supports advanced features for analysis, requirement and process services. EA modeling capabilities of business processes allows visualize business processes by UML profile for BPMN, to generate executable BPEL scripts from BPMN models and to verify the correctness of BPMN models.

The patterns are an excellent tool for building adaptive systems and reusable components. When in the new software development a repeating model is detected, this model, which is created in EA at the prior developments, may be applied, as a pattern. This is done through appropriate settings, for example, of the variables that allow integration with current development. The pattern provides a complex decision of abstract problem. That is why after the developer has stopped an appropriate pattern, it is necessary to modify it so as to meet the needs of current development. EA allows use of different patterns. Before this, they need to create, as such patterns [12]. For this purpose designers developed a standard UML [8] diagram and it is saved as a pattern in XML format. Thus generated XML file could be imported, as a resource to be used in the new system model development.

Another pattern feature is that it describes the basic concept which offers a suitable solution for a particular practical problem and a common development approach. That's why patterns can be adapted and reused in similar situations. They are defined independent of the application and also include a steps sequence description necessary to incorporate them as an integrated element of the developing system. EA patterns are basic building blocks of the conceptual data model. The patterns themselves are organized independently and reflect the proven skills and practices, by the scientific and practical viewpoint. They can easily be applied by a wide range of users, such as CASE environment allows improvement of the existing patterns, regardless of the specific particular design limitations. Moreover, patterns can be combined from several collections that reflect different classifications or architectural

layers. Thus the combination may lead to an improvement of the existing or introducing new patterns.

Fig. 4. represents the ideal case in which the model is developed, by the patterns. More than one pattern collection can be used for the specific development. Developers with the help of potential users, choose patterns suitable for solving problems in a particular organization for which the system was developed. To enable patterns to be jointly used, they need first to be determined the type of relationship between them. Then by UML diagrams, the patterns integration is realized. After the pattern integration thus has created a coherent conceptual model, the application, which meeting the organization requirements, is obtained. The conceptual model can be used as a basis for a new architectural pattern development.

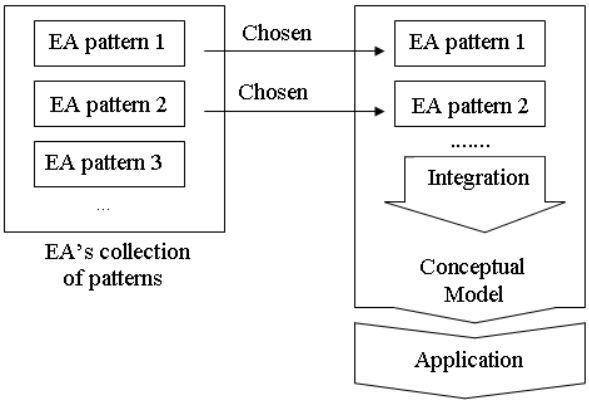


Fig. 4. The EA approach based on the use of patterns.

The same sequence of actions is using in order to integrate the patterns as the integration of the components of a software system. A simple way to integrate two patterns is detection of one or more identical classes in them. Identical classes are used as a point of integration. This way is typical uses relationship between patterns. It is necessary to determine which one of the two patterns is with a global application in the system. It will use the identical class of the local patterns. This is possible, because of the description and documentation of EA patterns, UML is used. Another approach to integration [14] involves the introduction of new relationship between the two classes of different patterns. Potentially this can be done with inheritance, when the relationship between patterns is refine. For example, class "Business Application" from one pattern can inherit the life cycle attributes of the class "Element of the life cycle" of the other pattern.

The main purpose of EA environment is supporting the development of business application, which is facilitated by the ability to use a wide variety of prepared, documented and proven practical patterns. For example, the classic patterns can be accessed by multiple toolbars. They are supported by the MDG (Model Driven Generation) technology. It extends the EA capabilities, by providing additional tools, patterns, UML configurations and other modeling resources. The elements of the classic patterns are in file EABase.eap, which is opened in the resources hierarchy window. These patterns can be inserted in the development project through three toolbars for creating, behavioral and structural patterns or through the context menu. When menu option More tools|GoF Patterns is selected, the toolbar that contains all 23 classic patterns, opens. We can generalize that to integrate a pattern in the system model, first it must be inserted into the model. So it becomes available from the resource box and optionally from the toolbar. In addition to existing patterns at EA new own

pattern can be created [13]. So first the pattern should be developed as a standard UML class diagram.

In order to ensure the better management and reliability, only users with appropriate permissions can record a UML diagram as a pattern and create it, when working with Ultimate editions of EA.

The patterns role is found on several levels depending on their purpose. First are the patterns that are used in the development of a specific code. They are the patterns from the lowest level, which are directly dependent on the requirements of a particular language or programming environment. They are specific and difficult can transfer to other software tools [6]. They are followed by patterns for the design of a particular component. This is the level at which most often patterns are applied. The patterns are applied at the level of architecture design. Using a pattern is not strictly defined and formalized. The different architectural models are not presented with specific patterns that can be directly invested in the development. For a description of the specific problem it is best to apply the design language, such as UML. An advantage of the EA is that it supports different levels of patterns.

Conclusion

Here is presented a possible design patterns classification, which includes three primary relationships between patterns and a variety of secondary relationships between patterns. This classification allows analyse existing patterns, organise them and to create new patterns. Relationships between existing patterns can be described by classification. The patterns themselves can then be catalogued or classified based on their relationships to other patterns. If the relationship of a new pattern to other patterns is described by classification, it should be easier to understand the pattern's relationships.

The patterns are a powerful tool for giving experience and ensure the quality of software, but tight deadlines for development applications many times lead to inappropriate pattern use. The classifications problem is that the methods and the results of the patterns relationships are not fully described. The main patterns advantage lies in the possibility to cover systematically repeated solution of a problem and to describe its applicability. Therefore patterns description language is the most appropriate tool for the realization of this dignity. CASE environment EA is extremely flexible and powerful development tool because it supports both different types of patterns and method UML.

References:

1. Alpert, S. R., Brown, K., Woolf, B., The Design, Patterns Smalltalk Companion, Addison-Wesley, 1988.
2. Bernus, Peter: Enterprise models for enterprise architecture and ISO 9000:2000. In: Annual Reviews in Control 27 (2003), pp. 211-220.
3. Braun, C., Winter, R., A Comprehensive Enterprise Architecture Metamodel and Its Implementation Using a Metamodeling Platform. EMISA 2005, pp. 64 - 79.
4. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., and Stal. M.: Pattern-Oriented Software Architecture, John Wiley & Sons, 1996.
5. Coplien, James O., Multi-Paradigm Design for C++, October 1998.
6. Coplien, James O., Schmidt, Douglas C., Pattern Languages of Program Design. Addison-Wesley, 1996.
7. Dyson, P., Patterns for Abstract Design. PhD thesis, University of Essex, 1997.
8. Filipova, N., Capabilities for Object oriented method UML in informational business modeling, UE-Varna almanac, 2005, 216-261c. ISSN: 08616752.
9. Filipova, N., Filipov, F., Guidelines for converting UML business models, UE-Varna news, 2010, 31-45c. ISSN: 13100343.
10. Fowler M. (2007), Patterns of Enterprise Application Architecture. Addision Wesley, ISBN: 0321513754.
11. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns, Elements of Reusable Object-oriented Software, Addison Wesley Professional, Boston (1995).
12. Gur, N., Steps to create enterprise/system Architecture http://weblogs.asp.net/ ngur/articles/194704.aspx 2004.
13. Harmon, P., Developing enterprise architecture, white paper, business trends., 2003.
14. Hoogervorst, J.A.P., Enterprise Architecture: Enabling Integration, Agility and Change, 2003.
15. Lorenz, D. H., Tiling design patterns — a case study. In ECOOP Proceedings, 1997.
16. Martin, R. C., Riehle, D., Buschmann, F., Pattern Languages of Program Design, volume 3., Addison-Wesley, 1998.
17. Riehle, D., Composite design patterns. In ECOOP Proceedings, 1997.
18. Tichy, W. F., A catalogue of general-purpose software design patterns. In TOOLS USA 1997, 1997.
19. Vlissides, J. M. , Coplien, J. O., Kerth, N. L., Pattern Languages of Program Design, volume, Addison-Wesley, 1996.
20. Zimmer, W., Relationships between design patterns. In Pattern Languages of Program Design, Addison-Wesley, 1994.

Smart Homes Cyberthreats Identification Based on Interactive Training

Zlatogor Minchev[1], Luben Boyanov[2]

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
Acad. Georgi Bonchev Str., Bl. 25A, zlatogor@bas.bg[1], lb@acad.bg[2]

Abstract. The paper briefly describes cyber threats identification framework in a smart home test bed environment. A problem space is initially built, through a q-based survey of potential cyber threats sources. Further on, an expert based reference with morphological and system modelling and analysis of this space is performed via a specialized software environment – I-SCIP-MA-SA. The validation is organized through an interactive human-machine agent based constructive simulation. Selected users' and environment characteristics are monitored for the proposed identified threats validation. This assures an explanatory cyber threats identification combined with real test bed experiments through interactive training.

Key words. cyber threats, smart homes, morphological and system analysis, agent based simulation, constructive simulation, interactive training.

1. Introduction

Today smart homes are becoming an indispensible modern part of our everyday live. This concept has gone through a significant evolution for almost a century and nowadays is addressing telemedicine, security and emergency areas, green energy and emerging technologies. This is also a result of the current fast ICT progress and opens a number of threats for the technologies development perspectives and their users' response.

The modern IT world is a place where the interactiveness between technologies and humans' is constantly evolving, mixing physical and digital realities. Humans are becoming more and more significant, with special attention paid to the behaviour dynamics and ambient factors influence. This could be a useful information source in the analysis of cyber threats for smart homes.

The paper briefly considers a methodological framework for identifying cyber threats on the basis of experts' believes, filtered with morphological and system analysis. The results are further experimentally validated in a test bed environment for a selected scenario sets. The process is organized around a multi-agent constructive simulation concept with human-in-the-loop participation. During this, the activities of the test bed inhabitants (smart homes users) and environment conditions are monitored for experimental correlation of selected scenarios sets.

This useful combination is producing a promising base for empirical validation of experts' believes through interactive training and constructive simulation. Further on, the paper gives a more detailed description of the proposed idea.

2. Methodological Framework

Generally, the methodological framework for cyber threats identification (see Figure 1) encompasses the application of the "scenario method" combined with analysis and

validation. Both implementing experts' believes and real human-in-the-loop monitored participation through agent based constructive simulation.

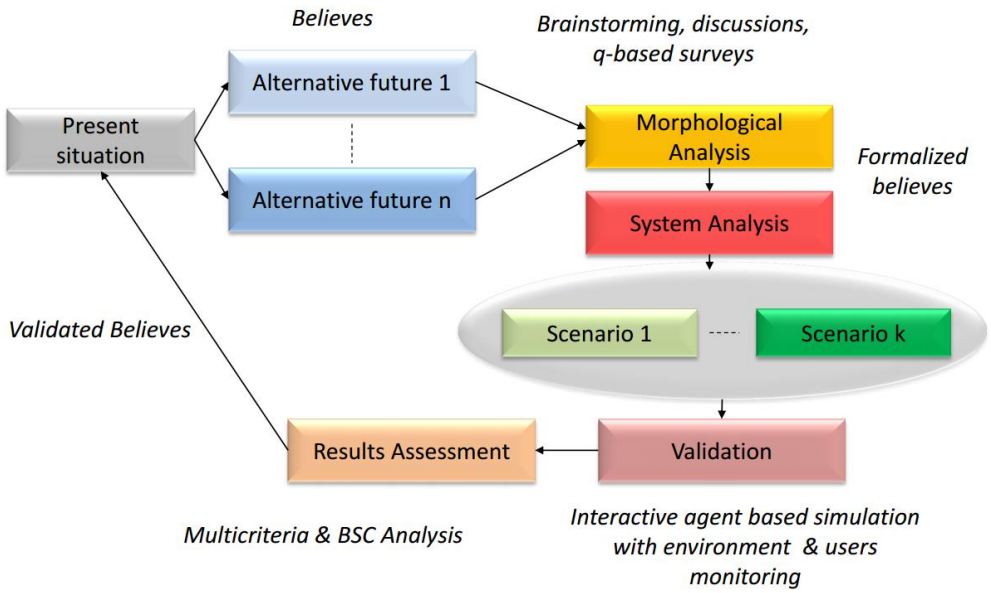


Figure 1. General schematic representation of the methodological framework.

The methodological framework, presented on Figure 1 could be practically implemented for smart homes cyber threats identification in four stages: (i) cyber threats identification, (ii) context definition, (iii) analysis and (iv) validation [11]. The next section gives more details of this four stages practical implementation.

3. Practical Implementation

Cyber threats identification was performed collecting users' focus group data gathered from two q-based surveys.

The first survey concerned Web technologies trends for a five years' time horizon and expected cyber threats in several social facets (*Civil Society*, *Banks and Finances*, *State Governance*, *Critical Infrastructure*, *Emerging Technologies*, *Education*). A focus group of 150 participants (national and international experts) has been studied [11].

The second survey results were produced from 250 participants' focus group at the University of National & World Economy – Sofia, The College of Telecommunications and Post and VISENSI Ltd. The obtained trends were covering: "*Type of used smart devices"*; "*Activities for using smart devices*"; "*Positives of using smart devices with Internet access"*; "*Negatives of using smart devices with Internet access"*; "*Type of information exchanged via smart devices*"; "*Smart devices influence to everyday life*".

A generalization of the obtained results from both surveys is presented in Figure 2 and Figure 3.

Technology/Dimension	Civil society	Banks & finances	State governance	Critical Infrastructure	Emerging technologies	Education
Web 1.0						
Web 2.0 / Web 3.0						
Web 4.0						
Web 5.0						

Figure 2. Q-based survey results about Web technologies trends amongst 150 participants.

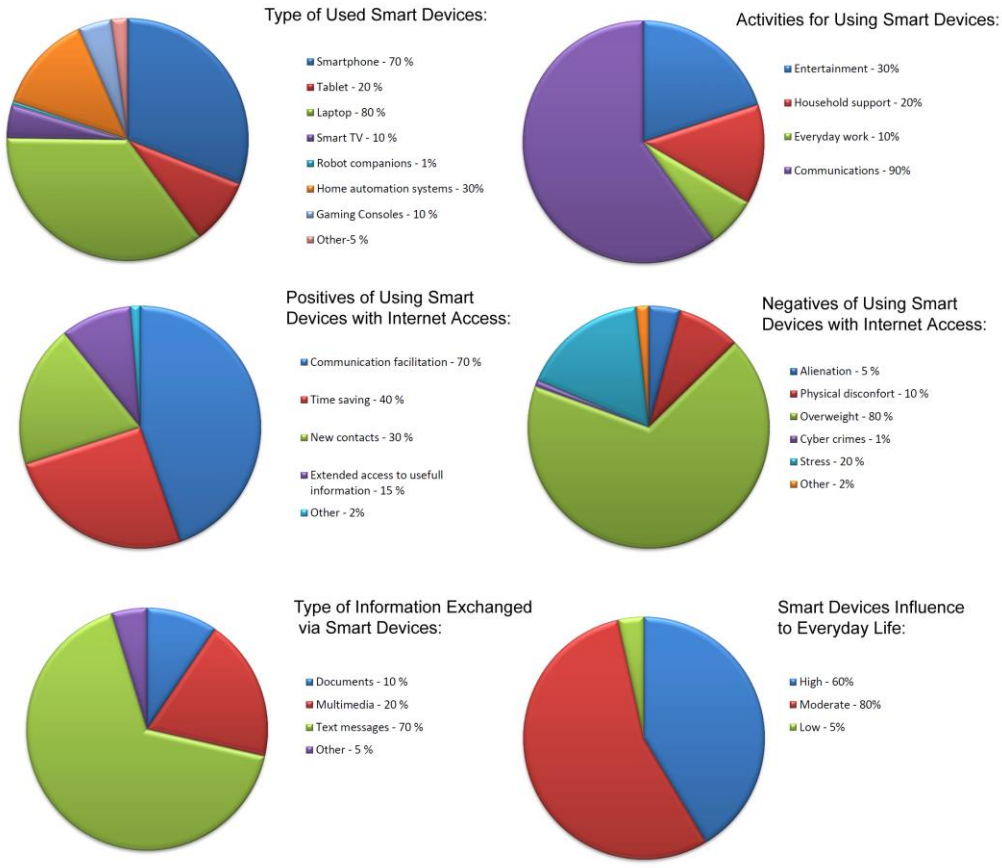


Figure 3. Q-based survey generalized results of smart devices amongst 250 participants.

The notations from Figure 2 [11] are using a discrete five-level color scale from "green" to "red" trough "yellow" that shows an increasing influence towards red and a decreasing one – towards green. The "blue" color is noting uncertainty. Similarly to another recent EU study, the selected time horizon was five years [1].

Briefly, the resulting trends are marking cyber risks and threats' importance increase in all Web 1.0/Web 5.0 technological areas. A visible exception of the part concerning Web 4.0/Web 5.0 (for *Banks and Finances* and *Emerging Technologies* facets) is quite understandable as these new technologies are expected to be available in at least ten-year time horizon.

The generalized results from Figure 3 demonstrate that most of studied the users are relying on *Smartphones* (70%) and *Laptops* (80%) for *Communication* (90%) and *Entertainment* (30%). A few of them are using smart devices for *House hold support* (30%), including: automated washing machines, drying machines, cleaning and cooking robots, dishwashers. Only 10 % from the users are having *Smart TVs* and *Gamming consoles*.

According to the studied users' focus group, the positives of smart devices usage are related to: *Communication facilitation* (70%) and *Time saving* (40%). The creation of *New contacts* is 30% of the users' priorities. The negatives are given to *Overweight* (80%) and *Stress* (20%). Mobile smart devices are most often used for text messages exchange (70%) and multimedia (20%). Currently the influence of smart devices influence to our everyday lives is classified as *Moderate* (80%) up to *High* (60%). The percentages sum is over a hundred, as the participants were allowed to give more than one answer.

Context definition is the second stage following cyber threats identification. Graphically, it is summarized in Figure 4:

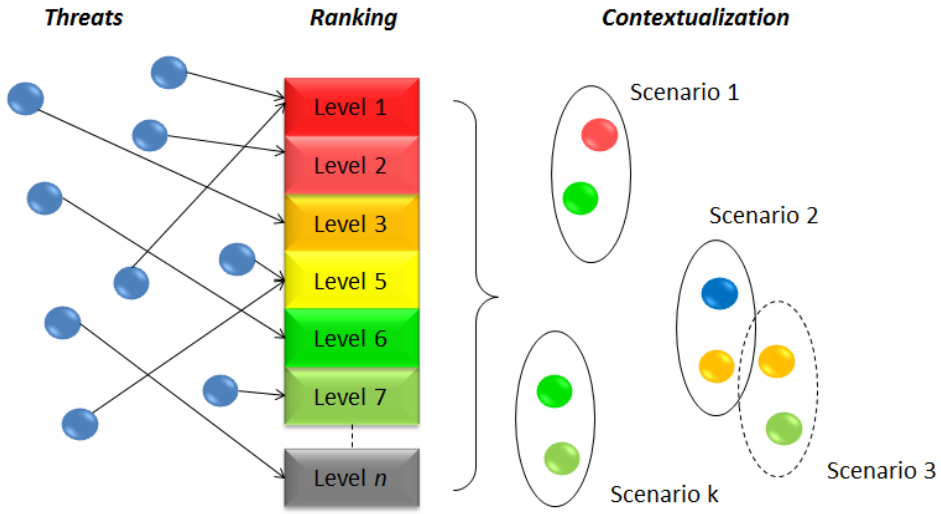


Figure 4. Graphical interpretation of the context definition process.

The identified cyber threats are arranged in accordance with their importance using experts' opinion. The results are implemented as driving factors [13] for the morphological and system analysis during scenarios preparation (see next stage). A good comprehensive example is the recent SysSec consortium study [7]. The authors mark three basic directions: "mobile devices", "social networks", "critical infrastructure".

A classification regarding smart homes environment, presented in Table 1 was proposed by the authors [2].

SMART HOME SERVICES	POSSIBLE THREATS	CRITICAL ATTACK POINTS	POSSIBLE CONSEQUENCES FROM THE ATTACK
Health care	Do not take medicine, pacemaker malfunctioning, etc.	Sensors, video surveillance, communication system, integrating system, external communications	Critical
Care for children or people with disabilities	Requires attention	Sensors, video surveillance, communication system, integrating system, external communications	Critical
Security and safety	Intrusion	Sensors, video surveillance, communication system, integrating system, external communications	Critical
Care for children or people with disabilities	Requires attention	Sensors, video surveillance, communication system, integrating system, external communications	Critical
Home environment	Fire, flooding, gas leakage	Sensors, video surveillance, communication system, integrating system, external communications	Critical
Smart home appliance	Does not turn off, turns on/off at wrong time	Sensors, video surveillance, communication system, integrating system	Non-critical, but dangerous
Privacy	Violation of privacy, data gathering	Video surveillance, communication system, integrating system, external communications	Non-critical but dangerous
Entertainment and pleasure	Malfunctioning of the pleasure, comfort and entertainment systems	Sensors, communication system, integrating system	Non-critical

Table 1. Services, dangers, attack points and consequences in smart homes.

The results from Table1 are practically used during the analysis and validation methodological stage.

The **Analysis** is performed in four sub stages. As far as the scenario method has been selected for methodological base, morphological and system analyses are initially performed. Generally, these processes are relying on experts' data with a lot of combinations and uncertainties, so some software support with I-SCIP-MA-SA [8] is also implemented.

The key idea for machine interpretation of the problem space is the E-R paradigm [3] graphically represented with named round rectangles (Entities) and weighted headed arrows (Relations).

Morphological analysis

A resulting problem space for smart homes cyber threats scenarios exploration through morphological analysis with 16 alternatives spread in 5 dimensions [9] is shown on Figure 5. The dimensions are presented in different colors.

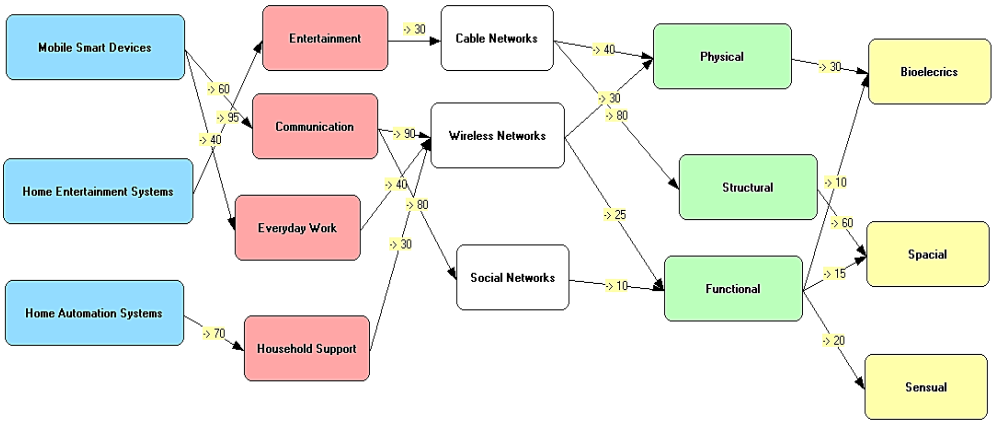


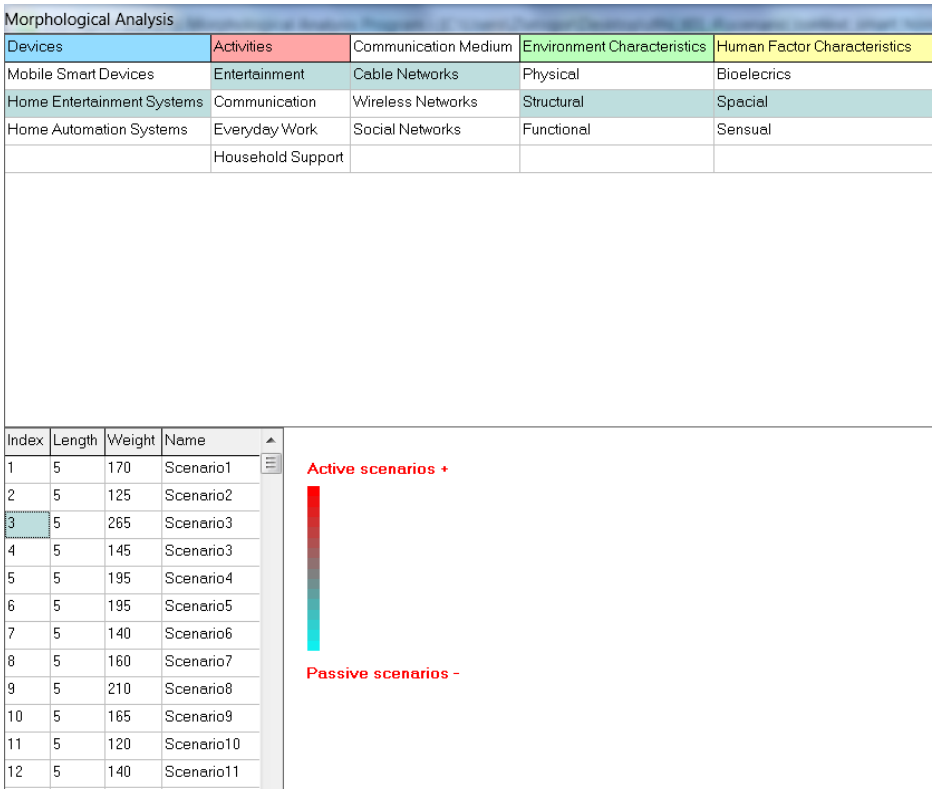
Figure 5. A screen shot of I-SCIP-MA for morphological scenario problem space with 16 alternatives and 5 dimensions, studying smart homes cyber threats.

The resulting cross-consistency matrix contains $N = 1620$ ($N = 5 \times 3 \times 4 \times 3 \times 3 \times 3$) scenario combinations (see Figure 6). It is important to note here, that the implemented scale for weighting the arrows gives the percentage measure, covering three levels: weak [0-30%], moderate [30-50%], high [50-10%] from the interval [0, 1].

The *Dimensions* part is encompassing the following: "Devices", "Activities", "Communication Medium", "Environment Characteristics", "Human Factor Characteristics". Each dimension contains different number of alternatives that in practice are subspaces, e.g. "Devices" encompasses: "Mobile Smart Devices", "Home Entertainment Systems" and "Home Automation Systems".

A ranking, using Relative Common Weight (RCW) has been performed. RCW sums the unidirectional relations' weights (noted with yellow labels above the relations) connecting an alternative from each of the five dimensions that were used. The final results are scenarios with negative or positive RCW in accordance with obvious or hidden cyber threats identification.

The most interesting scenario combinations in our morphological analysis were: Scenario 3 (RCW = 265, encompassing: "Home Entertainment Systems" → "Entertainment" → "Cable Networks" → "Structural" environment characteristics → "Spacial" human factor characteristics); Scenario 9 (RCW = 210, encompassing: "Mobile Smart Devices" → "Communication" → "Wireless Networks" → "Physical" environment characteristics → "Bioelectric" human factor characteristics) and Scenario 19 (RCW = 110, encompassing: "Mobile Smart Devices" → "Everyday Work" → "Wireless Networks" → "Functional" environment characteristics → "Bioelectric" human factor characteristics).



Morphological Analysis

Devices	Activities	Communication Medium	Environment Characteristics	Human Factor Characteristics
Mobile Smart Devices	Entertainment	Cable Networks	Physical	Bioelecrics
Home Entertainment Systems	Communication	Wireless Networks	Structural	Spacial
Home Automation Systems	Everyday Work	Social Networks	Functional	Sensual
	Household Support			

Index	Length	Weight	Name
1	5	170	Scenario1
2	5	125	Scenario2
3	5	265	Scenario3
4	5	145	Scenario3
5	5	195	Scenario4
6	5	195	Scenario5
7	5	140	Scenario6
8	5	160	Scenario7
9	5	210	Scenario8
10	5	165	Scenario9
11	5	120	Scenario10
12	5	140	Scenario11

Figure 6. A screen shot from I-SCIP-MA of morphological cross-consistency matrix with $N = 1620$ scenario combinations for smart homes cyber threats exploration.

These results do not give any concrete cyber threats identification scenarios but just outline, in accordance with the experts believes, the importance of the activities "Entertainment", "Communication" and "Everyday Work" together with the human factor "Spacial" and "Bioelectric" characteristics in smart home environment.

System Analysis

Generally, the system analysis for a smart home is a complex dynamic system approximation. The system model for smart homes could be both static [2] and dynamic. Whilst, the static one gives general classification of the objects of interest, the dynamic is quite useful in the *Validation* process. It is important to note that the E-R paradigm is used implementing weights (with similar to the morphological analysis scale) and time of the bidirectional relations (both noted, consecutively with labels in yellow and blue).

Additionally, a three dimensional Sensitivity Diagram (SD) presenting influence (x), dependence, (y) and sensitivity (z). Four sectors encompassing the entities classification are utilized: green (buffering), red (active), blue (passive) and yellow (critical). All entities from the model are visualized in SD with indexed balls.

Practical illustration of a static smart home general system model and a resulting SD classification are given in Figure 7.

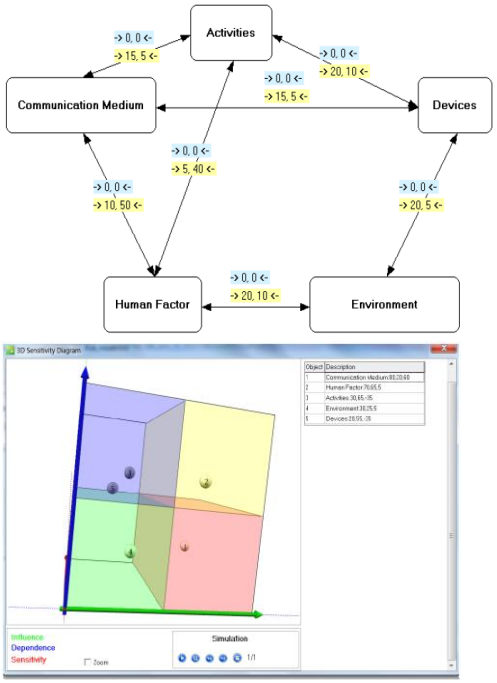


Figure 7. A smart home system model (left) and resulting SD diagram (right) after [2].

The resulting SD from Figure 7 gives a profitable classification for further analysis, outlining the "Human Factor" (indexed ball/sphere "2" with coordinates {x=70, y=65, z=5}) as a critical entity together with the potential hidden cyber threats passive entities: 'Devices'(indexed ball "5" with coordinates {x=20, y=55, z = -35}), "Activities" (indexed ball "3" with coordinates {x=30, y=65, z= -35}) and real active one: "Communication Medium" (indexed ball "1" with coordinates {x=80, y=20, z=60}).

Further *Validation* was performed trying to obtain more comprehensive evidence for this experts' believes and analyses results.

Validation

This final stage of the presented methodological framework is related to interactive training and includes agent-based modeling and simulation. The presented multi-agent model, implemented in I-SCIP-SA dynamic environment (see Figure 8) is encompassing seven key role agents: "Real Human Agent", "Entertaining Agent", "Comms Agent", "Storing Agent", "Monitoring Agent", "Digital Assistant", "Attack Agent". Generally these roles cover a number of devices, protocols and parameters. What is important to note here is the capability for connecting real devices with simulated ones, i.e. mixing the virtual and real world. Additionally, Figure 8 illustrates the initial believs evolution, starting from green and progressing towards different SD sectors in accordance with model relations weights' dynamics.

As the human factor in this dynamic model (noted with indexed "2" sphere) is classified as a passive one (blue sector of SD on Figure 8) a more detailed analisys could be performed trough a test bed interactive environment.

For this purpose, a smart home test bed environment (see Figure 9) has been organized in the framework of DFNI T01/4 project at the Institute of Information and Communication Technologies (IICT), Bulgarian Academy of Sciences – BAS.

The test bed environment is positioned in a room equipped with a number of smart devices, including: 3D TV/monitors, X-box game console, entertainment and cleaning robots, programmed tablet remote control and IP video omnidirectional monitoring system. An ad-hoc created digital assistant - "Alex", provides voice control for lighting, multimedia and heating with holo-like projection avatar. In addition, an environment embedded Xbee sensor barometer system and wearable human factor bio headband are being developed [4]. The sensor barometer system is also capable to monitor CO/CO2 concentration measurement, radiation, electromagnetic fields and dust particles [5].

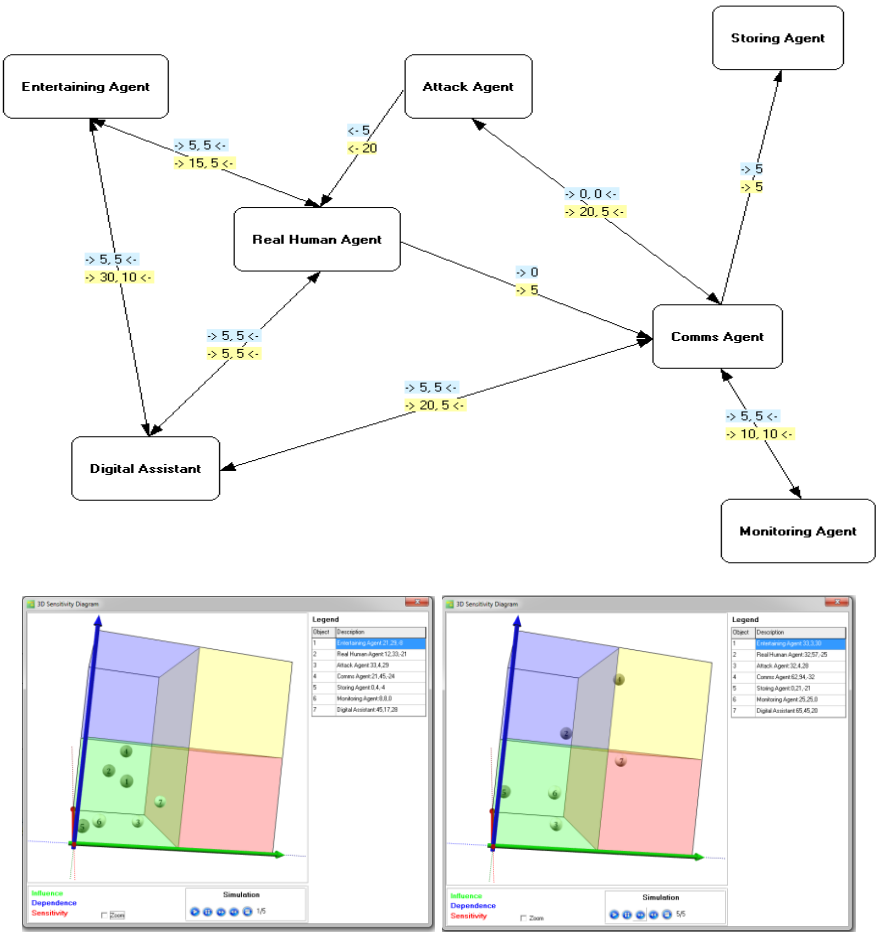


Figure 8. Basic multi-agent system model representation with 5 steps dynamics.

The human factor activities are monitored via EEG Nation 7128W – C20, bio headband for ECG and body temperature monitoring [4]. All data from sensors and video behavior monitoring is stored in a data base.

Figure 9. Moments of smart home test bed validation usage at IICT-BAS.

The above described test bed has been organized for practical agent-based interactive training.

Some interesting results have been recently published [12] concerning the human factor hidden threats related to EEG dynamics in 2D and 3D visual environment and the gamming process in social networks. The results directly accentuate the necessity of comprehensive study of cyber threats problem in smart homes and the emerging "digital drugs" [6]. We are planning further improvement of the validation process with other users' behavioural modalities. The results will be evaluated via multicriteria analysis and balanced score card [10].

Discussion

The progress in smart homes technologies is opening a number of cyber threats to their users today. Whilst some of them are quite obvious, other related to entertainment, privacy and appliances are hiding a lot of unexplored domains. Examples for such new cyber threat areas are the digital drugs (addiction to technologies) and social engineering that are important for the future generations of inhabitants (users) of smart homes.

A suitable framework approach for studying these problems is the combination of experts' believes data, analysis, modelling, inhabitants and environment monitoring, as well as, practical validation through real constructive experimental training.

Acknowledgement

The results in this publication are financially supported by: "A Feasibility Study on Cyber Threats Identification and their Relationship with Users' Behavioural Dynamics in Future Smart Homes", Research Grant "Funding of Fundamental & Applied Scientific Research in Priority Fields", Bulgarian Science Fund, Ministry of Education Youth and Science, 2012-2014, DFNI-T01/4, www.smarthomesbg.com. A special gratitude for the threats identification methodology is also given to: "EU Network of Excellence in Managing Threats and Vulnerabilities for the Future Internet" – SysSec, FP7 Grant Agreement No. 257007, 2010 – 2014, www.syssec-project.eu

References

1. Balzarotti, D. Second Report on Threats on the Future Internet and Research Roadmap, SysSec Consortium Deliverable D4.2, September, 2012, http://www.syssec-project.eu/m/page-media/3/syssec-d4.2-future-threats-roadmap-2012.pdf [Online]

2. Boyanov, L., Minchev, Z. Cyber security Challenges in Smart Homes, In Proceedings of NATO ARW "Best Practices and Innovative Approaches to Develop Cyber Security and Resiliency Policy Framework", Ohrid, Macedonia, June 10-12, 2013 (in press)
3. Chen, P. The Entity-Relationship Model-Toward a Unified View of Data, ACM Transactions on Database Systems, 1, no.1, 9–36, 1976.
4. Georgiev, S., Minchev, Z. An Evolutionary Prototyping for Smart Home Inhabitants Wearable Biomonitoring', In Proceedings of Conjoint Scientific Seminar "Modelling & Control of Information Processes", IMI-BAS, 21 – 31,`November 19, 2013.
5. Georgiev, S., Kolev, H., Obreshkov, N., Lalev, E. Security System for Future Smart Homes, In Proceedings of National conference with international participation in realization of the EU project "Development of Tools Needed to Coordinate Inter-sectorial Power and Transport CIP Activities at a Situation of Multilateral Terrorist Threat. Increase of the Capacity of Key CIP Objects in Bulgaria", at Grand Hotel "Sofia", Sofia city, Bulgaria, June 4, 91-100, 2013 (in Bulgarian).
6. Guma, G. Messing with Our Minds: Psychiatric Drugs, Cyberspace and "Digital Indoctrination", Global Research, November 11, 2013,http://www.globalresearch.ca/messing-with-our-minds-psychiatric-drugs-cyberspace-and-digital-indoctrination/5357710 [Online]
7. Markatos, E., and Balzarotti, D. (Eds) The Red Book. The SysSec Roadmap for Systems Security Research, The SysSec Consortium, 2013, http://www.red-book.eu/m/documents/syssec_red_book.pdf [Online]
8. Minchev, Z., Shalamanov, V., Scenario Generation and Assessment Framework Solution in Support of the Comprehensive Approach, In Proceedings of SAS-081 Symposium on "Analytical Support to Defence Transformation", RTO-MP-SAS-081, Sofia, Boyana, April 26 – 28, 22-1 – 22-16, 2010, http://www.gcmarshall.bg/KP/new/MP-SAS-081-22-MINCHEV-SHALAMANOV.pdf [Online]
9. Minchev, Z., Boyanov, L., & Georgiev, S. Security of Future Smart Homes. Cyber-Physical Threats Identification Perspectives, In Proceedings of National conference with international participation in realization of the EU project 'Development of Tools Needed to Coordinate Inter-sectorial Power and Transport CIP Activities at a Situation of Multilateral Terrorist Threat. Increase of the Capacity of Key CIP Objects in Bulgaria', Grand Hotel 'Sofia', Sofia, Bulgaria, 165-169, June 4, 2013.
10. Minchev et al., Joint Training Simulation and Analysis Center, Technical Report, Institute for Parallel Processing, Bulgarian Academy of Sciences, 2009, http://www.gcmarshall.bg/KP/new/jtsac_tr.pdf [Online].
11. Minchev, Z. Security of Digital Society. Technological Perspectives and Challenges, In Proceedings of Jubilee International Scientific Conference "Ten Years Security Education in New Bulgarian University: Position and Perspectives for the Education in a Dynamic and Hardly Predicted Environment" Sofia, Planeta -3 Publishing House, 438-444, 2013 (in Bulgarian)
12. Minchev, Z. 2D vs 3D Visualization & Social Networks Entertainment Games. A Human Factor Response Case Study, In Proceedings of 12th International Conference, ICEC 2013, São Paulo, Brazil, October 16-18, 2013 (Editors: Junia C. Anacleto, Esteban W. G. Clua, Flavio S. Correa da Silva, Sidney Fels, Hyun S. Yang), Lecture Notes in Computer Science, vol. 8215, 107-113, 2013b.
13. Schwartz, P. The Art of the Long View: Planning for the Future in an Uncertain World, Doubleday Currency, New York, 1991.

"Scientific information has the power to transform our lives for the better - it is too valuable to be locked away."
Neelie Kroes
European Union Commissioner for the Information Society

Academic Use of Open Digital Archives

Juliana Peneva, Stanislav Ivanov [1]

[1] NBU, Dept.of Informatics
21 Montevideo St, 1618 Sofia, Bulgaria
jpeneva@nbu.bg; sivanov@nbu.bg

Abstract. Open access academic archives are the proper instrument to make visible the majority of the coursework done during the years and to propagate the results of already paid research activities, thus raising their value. By offering proper services, archives become the media which facilitates the forming of interim research and learning societies among students and teachers. This is the way to achieve an improved management of intellectual assets of the whole university community. In this paper the authors try to summarize the advantages of open access university archives and to discuss some issues arising from rules for deploying and accessing course materials and research papers.

ACM Classification Keywords. H3.5 On-line information services– Data sharing; H3.7 Digital libraries-Collection.

1. Introduction

Universities follow different policies when exposing their intellectual assets. Some of these institutions deliver open courseware, whilst others prefer to represent only their organization and facilitate the communication at their sites. Meanwhile, digital repositories have become significant though controversial means of communication among researchers and lecturers. Institutional digital archives are a natural environment for the deployment and public (even internal) assessment of various materials related to the teaching and research activities of the community members. Members of educational institutions can use the repository in many different ways. Here again, the local policy may facilitate and/or restrict these activities. Universities tend to deploy and exhibit different kinds of their intellectual assets such as learning content and research papers. It is a matter not only of user's convenience, but of representativeness and prestige as well. The rapid advance of information technologies made possible not only to manage large institutional archives and to equip them with various services, but to organize similar facilities at lower, departmental level.

During the last decade different types of repositories ranging from digital libraries through various institutional collections and e-journals up to collaborative learning environments have been built. Not surprisingly the main share of active repositories belongs to countries with advanced higher education and science. Members of educational institutions and research centers deliver and use the repository content intensively. The main

reason for their continuous activity delineates different purposes ranging from free access to deposited resources through using a variety of services.

Digital repositories for academic purposes mark a continuous development. In this perspective, universities and scientific institutions demonstrate a remarkable activity. Since the beginning of year 2007 the growth of such repositories listed in the OpenDOAR Database [8] shows a constant increase of 100 repositories per year up to its present number of over 2500. Currently OpenDOAR report over 2500 storage facilities. For the preceding calendar year, even there was an increase of about 300. There is a sustained interest from universities to the use of open digital archives. In Bulgaria there are registered six digital archives of open access: two at the Institute of Mathematics and Informatics; on at Medical University of Sofia, Bourgas Free University, Sofia University and New Bulgarian University respectively.

New Bulgarian University was among the first academic institutions in the country that developed and maintains an open digital archive. The archive represents a cardinal digital environment for long-term preservation of the electronic scientific output of NBU academic staff and post-graduate students. Currently New Bulgarian University Scholar Electronic Repository contains 1275 items consisting of the research output of the institution. Users may set up Atom and RSS feeds to be alerted to new content. The interface is in English. The subjects are multidisciplinary; the content comprises articles, conferences, books and learning objects [6].

Apart from this, but using the same technology, a departmental repository was developed to assist all non-auditoria activities. The main goal is to manage and disseminate digital materials created by the department and its community members [9]. The repository will be used for electronic publishing and housing of different digitized collections concerning the knowledge resources of the department. The final goal is to offer open but local access to scholarly research. So, the departmental repository is designed to deploy content not covered by the university infrastructure: investigations, learning resources, theses, students' projects and papers.

Some issues arising from rules for deploying and accessing course materials and research papers deserve close attention. Doubtless it is better to make visible the majority of the coursework done during the years, but copyrights should be respected and regulated properly. Open access archives are the suitable instrument to propagate the results of already paid research activities, thus raising their value. However the research teams should be informed about this possibility and encouraged to use it. Also, having appropriate services, archives would be the media which facilitates the formation of interim research societies involving students in them. This is the way to attain an improved management of intellectual assets of the whole university community. Digital archives have great potential for any kind of value added services as well.

In the context of the above, the main goal of this paper is to discuss the use of open digital archives. Taking into account our experience we present their benefits for community members, for the university and for the process of research generally. In Section 2 we present the key features of academic digital repositories that raise the representativeness of academia and intensify the scientific research. Section 3 deals with discussion of the local policy for management and internal use of repositories at subdivision level. We summarize our findings in Section 4.

2. Key features of academic digital repositories for the representativeness of academia scholar work

The successful management of digital resources is very important for any organization to realize a business advantage [1]. Not surprisingly they are considered as assets along with financial, material and human resources [2]. For higher education institutions digital resources can be used strategically to expose their intellectual assets such as learning content and research papers. Usually digital resources are organized as university-based institutional repositories, including long-term preservation and distribution thus permitting to derive their maximum business value. The collection of digital content into a repository enables higher education institutions to support research, teaching, learning, and administrative processes [3], [4].

According to the SPARC alliance [5] institutionally defined repositories are scholarly, cumulative, open and interoperable. The digital content is stored and managed to facilitate searching and retrieval of the collected items as well as their later reuse. The decision what to put into a repository depends on key institutional intentions and objectives. Some repositories store only particular items e.g. articles, books, works of art, etc. so long as others gather a significant amount of scholar work. The management of e- resources can be performed alternatively via virtual learning environments, wikis and other informal content sharing applications. However in our work we concentrate on digital archives capturing the intellectual product created by the stakeholders of the overall educational process: faculty, research staff and post-graduate students.

The repository of New Bulgarian University is governed by a supervising board called Evaluation Commission. The Head of Library and Information Services Department is acting as a repository editor thus enforcing compliance with certain rules when an item is going to be deposited. Academic staff and university PhD students are allowed to deposit their documents that might be not published or published via electronic or traditional means. In the case of published elsewhere documents the author who deposits have to hold the copy rights. If the copyrights belong to a publisher or other organization, a permission to deposit has to be granted. Academic staff can submit unpublished documents as well. In this case authors are expected to abide by ethical standards and to ensure quality content. The document types are listed in the user interface of the software used to implement the repository. E-Print [6] allows depositing articles, books, conference items, theses, artifacts, images, compositions, audios, videos etc.

The benefits of publishing in an institutional archive of open access can be summarized from different points of view:

1. For the university as an educational and research institution:
   - access to the intellectual output is provided

   The Scholar Electronic Repository represents the intellectual product created by the community members thus increasing the institution's visibility and its public value. The archive, by capturing and preserving collective intellectual capital, increases the overall institution's academic quality. There exists organizational support towards innovative means to research dissemination. In this way the sharing of ideas and know-how's as well as the rapid communication of research becomes feasible. Collaborative research is promoted. In addition an easy access to faculty papers is achieved. The demonstration of value can attract tangible benefits including project funding from both public and private sources. Documents are searchable via the Internet as they are indexed by search engines and made

accessible to a wider audience. In this way the production of the individual authors and the institution as a whole is promoted.

- enhanced scholar communication becomes possible

In the traditional system of scholarly communication much of the research findings are dispersed through different journals. However nowadays scholars use Internet intensively to disseminate their achievements. There is growth in the open access availability of research publications, both gold (author pays for publication) and green (self-archiving by the researcher). Approximately 30% of all articles are thought to be available as open access, two-thirds in green and one-third gold [7]. The university repository preserves and make accessible the staff intellectual output in a straightforward manner. In this way the foundation of a new lightened publishing model is set. Even articles published in academic journals can be placed in the repository to attain a larger audience. Global access to research literature is achieved. This is the way to complement and supplement journal publishing.

- sharing of learning and teaching materials locked in Virtual Learning Environments

Publication of educational resources in addition to the management of the curriculum is particularly important for dynamic areas of human knowledge. Provision of copyright curriculum in such areas is also a kind of advertising on campus. Making a learning content visible (open courses) increases the potential reuse of the materials.

2. For the academic staff

- improved research knowledge management

Sharing of research outputs, unpublished ideas and know-how facilitates their public visibility. In fact researchers manage and store digital content connected with their investigations including research data. Thus collaborative work on institutional projects is promoted. Community members have a place where their scholarly works are permanently exposed. The available scientific results give rise to the preparation of new joint projects. Knowledge sharing and reuse is facilitated.

- broad dissemination of published research findings

Because of the short period after uploading the document, authors are stimulated to publish thus achieving a fast dissemination of their results and a greater impact. In rapidly developing areas of scientific knowledge e.g. computer science or communications researchers can offer preprints via the open access archive in order to claim priority and to get a fast feedback.

- increased citation of some papers

Open access favors the dissemination of published research in the archives and contribute to the growth of citing of the scientific production. Citation analysis demonstrates that research papers that are freely available are easier to cite.

- indexing

Open access archives are indexed by search engines, which promote both the production of the individual authors and the institution as a whole. This allows for the creation of personalized publication lists and increases the citation. Via the usage of different metrics the researchers obtain hit rates on specific papers. The impact factor is also derivable.

Last but not least we have to stress some shortcomings. The quality of documents submitted to the open access archives are estimated by the supervisors following formal

criteria. The author's contributions might appear obsolete and/or disputable. Its author personal responsibility for the quality of content e.g. plagiarized texts, faults or junk.

3. Local management policy and internal use of academic digital repositories

In the dynamic domain of computer science it appears especially helpful to preserve some deliverables of the educational process itself like case studies, student's research projects, diploma theses, lecture notes and videos etc. It might be useful to deploy some intermediate results from various research activities, to announce ideas and proposals, seeking support, recognition and collaboration. In this way reports of research under development could be submitted to discussions and criticism among the members of a professional community. However similar writings are to be accessed thematically on a local basis. This is not to belittle such writings. Rather, they are generated as part of separate publishing activities. They might exhibit shortcomings, if any, not addressed by the present policies of an open access archive.

The goal of university research is the creation, dissemination, and preservation of knowledge. This is the way to disseminate good practices and tips among the learners. Sharing of the unpublished ideas and know-how's as well as a rapid communication among the researchers becomes feasible. Collaborative research is promoted. In addition an easy access to faculty papers by students is achieved.

Nevertheless that the Scholar Electronic Repository is announced as an open e- space some thematic subdivisions are to be with restricted (local) visibility. We propose local access to university projects in progress, the so called "gray literature", e.g. diploma theses, learning resources, students' projects, working papers, technical reports, presentations, etc. By depositing quality examples of students' paper work a significant support of students' endeavors is achieved and a location to deliver e-portfolios is submitted. Important students' works become easily available and good practices can be disseminated. Since no library can deliver all the resources students need, collecting quality examples of students' works creates a new layer of information that is readily locally accessible.

Publication of educational resources in addition to the management of learning content is particularly important for dynamic areas of human. In view of this, we have undertaken the creation of a departmental archive with main purpose to store additional teaching materials and well executed student works, assignments and theses. The departmental repository offers an efficient access to many useful deliverables of the educational process itself. It is well known that students in addition to learning and understanding existing knowledge need to produce new knowledge in order to be part of the knowledge society. So, the departmental repository by capturing and preserving collective intellectual capital becomes a vital component of e-learning and increases the overall institution's academic quality. The restricted access to files within the university limits the direct borrowing of texts. We believe that this is the way to warrant the correct usage of already published texts and the copyrights.

Considering the benefits to gain a restricted access to some deliverables of the educational and research activities it is up to the departments to take the necessary steps towards building of similar units of the university archive. We enjoy full maintenance from the university IT department and firm support from the superior administration.

4. Conclusion

Open digital archives are of great importance for the public visibility and recognition of universities. By allowing an improved management of intellectual outputs and freeing up the process of dissemination, their main purpose is to rise up the representativeness of the institution.

In this paper we discussed the academic use of open digital archives. We shared the reason for limited local access to some documents. An underlying level of locally visible target oriented units of the university archive is appropriate in order to support the local scientific exchange, interdisciplinary research and students' paper work, etc. Additional services to encourage community sharing and exchange of both practice and content could be developed.

The Scholar Electronic Repository is an open e- space which provides long-term preservation of electronic documents and assists in the dissemination of research findings both at public and local level.

References

5. Duncan C. Digital Object Repositories Explained, an Intrallect White Paper, October 2006.
6. Bluthe E., Chandra V. The Value Proposition in Institutional Repositories EDUCASE Review, September/October 2005.
7. Crow R. The Case for Institutional Repositories: A SPARC Position Pape SPARC, 2002.
8. http://www.jiscinfonet.ac.uk/infokits/digital-repositories/, accessed on January 09, 2014.
9. http://www.arl.org/sparc, accessed on January 09, 2014.
10. http://eprints.nbu.bg/, accessed on January 09, 2014.
11. Poynder, R. 2011. Open Access by Numbers, Open and Shut, http://poynder.blogspot.com/2011/06/ open-access-by-numbers.html, accessed on January 09, 2014.
12. www.opendoar.org, accessed on January 09, 2014.
13. Lynch C. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 (February 2003): 1-7.

DECISION CONTROL APPLICATION IN COMPLEX SYSTEMS: AN INOVATIVE DECISION-MAKING SUPPORT STRATEGY

Rumen Andreev, Yuri Pavlov
Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences,
E-mail: rumen@ isdip.bas.bg, yupavlov15@isdip.bas.bg

Abstract:

Decision-making support strategy is an important factor that determines a top-down approach to development of Decision-Making Support Systems (DMSS). This strategy depends on the objective that has to be achieved in supporting of decision maker. Since it is known four major objectives, we define the following strategies: problem formulation-driven, decision model-driven, decision-driven and decision implementation driven. Respectively, the paper presents a categorization of decision making support and an application of an innovative approach to decision making. It combines two strategies – model-driven and decision-driven.

The decision-driven DMSS concerns decision aiding that is considered in the context of Personal Decision Analysis (PDA). The paper presents a decision technology that is suitable for realization of PDA decision tool quantifying personnel judgment. This technology realizes measurement of human's, objectivly oriented preferences as utility function. The analytical description of the expert's preferences as value or utility function will allow mathematically the inclusion of the decision maker in the value based model description of the complex system "Technologist-process". In this way we introduce a Model-driven decision making strategy.

We demonstrate this system engineering value driven approach within determination of the equilibrium points in the competitive trading modeled by the Edgworth box. The contract curves and the trading equilibrium in the Edgworth box are specified on the individual consumers' preferences. The presented mathematical formulations serve as basis of tools development. These value evaluation leads to the development of preferences-based decision support in machine learning environments and iterative control design in complex problems.

Keywords: Decision Support System, Decision Making**,** Preferences, Utility, General Equilibrium, Edgeworth box.

Introduction

A great number of decision support systems (DSS) are produced with the objective to provide mechanisms to help decision makers get through a sequence of stages during decision making process in order to reduce uncertainly and processing of ambiguous decisions. In conformity with our viewpoint, the basis for defining DSS is determined by the perceptions of what a DSS does (supports decision-making in unstructured problems) and the objective(s) of this support – task(s), in which a decision maker has to be assisted during decision making process. That is why, in definition of these supporting systems we use the term decision making support system (DMSS) instead DSS.

In a top-down approach to development of DMSS, the determination of decision-making support strategy is very important. This strategy depends on the objective that has to be achieved in supporting of decision maker. In correspondence with the generic decision making process we can determine four major objectives of decision maker that he has to reach with the help of appropriate DMSS – construction of problem scheme, construction of decision model scheme, construction of decision scheme and decision implementation. Since

these objectives drive development of DMSSs, the corresponding DMSS come respectively under the following categories: *problem formulation–driven*, *decision model–driven*, *decision-driven* and *decision implementation-driven* (Fig. 1).

We divide implementation-driven DMSS in two classes, since it is possible to develop separately DMSSs for supporting task planning and task monitoring. Model-driven DMSS coincides with determined model-driven DSS. It is, by definition, functionally based on one or more quantitative models and is designed such that a user can manipulate model parameters in order to analyze decision alternatives. A DMSS may be classified as model-driven, if the model can be accessible to a non-technical user via an easy-to-use user interface, the model should provide a simplified, easily understandable representation of the decision situation, and the intention should be that the DSS itself is to be used repeatedly in the same or similar decision situation. Its main component is a model base that is a repository for the formal models of the decision problem and the methodology for developing results (simulations and solutions) using these formal models .

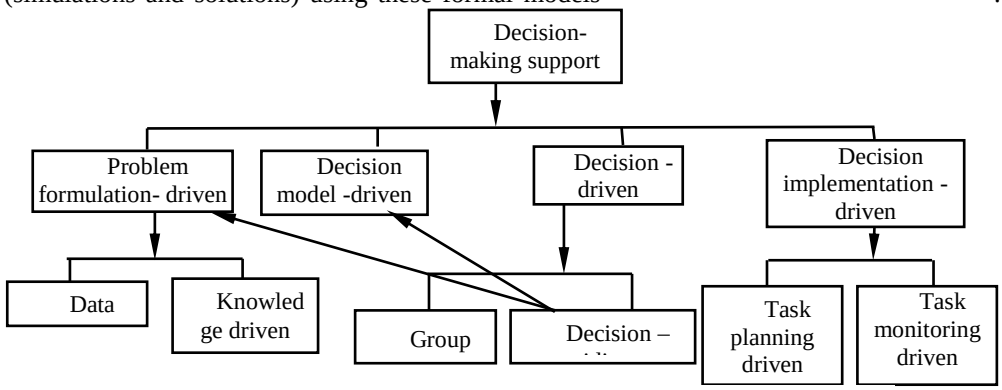


Figure 1. Categorization of decision making support

Decision-driven DMSS are divided into two classes – *decision aiding DSS* and *group DSS*. The development of decision aiding DSS is considered only in the context of personal decision analysis (PDA), quantifies judgment and processes it logically. The means used for quantify judgment are known as decision tools that realize certain decision technology. There is little "useful" research, in the sense of leading toward decision technology that advances the interests of the decider, through decisions that are either more sound or more effectively communicated. A decision technology has to (1) use all the knowledge a decider normally uses; (2) call for inputs that people can readily and accurately provide; (3) produce output that the decider can use (Baron, 2008). It is necessary to notice that, though recent research has made important scientific advances, little has influenced decision aiding practice, and little current research attacks the problems that are still holding back successful decision aiding.

The development of successful decision aiding tools requires more research in the following aspects of decision technology: decision strategy, principles of decision tools design and development of specific tools. The main objective of decision strategy has to be the best integration of informal analysis into informal reasoning without disrupting it. The principles of decision tool design are the following: how judgment-intensive should decision models be, in given circumstances; how well can people make hypothetical value judgments or hypothetical factual judgments. The development of specific tools concerns the next problems:

- Some mathematical algorithms call for inputs that fit people's cognition better than others. Inference by Bayesian updating, requires assessors to make hypothetical likelihood judgments. Are people better at making the resulting posterior assessments directly?
- What should be the form of utility elicitation – holistic, decomposed into additive pieces or further decomposed into factual and value judgments as in Multiattribute Utility Analysis.

Following Schmeidler people often wonder why economists analyze models whose assumptions are known to be critically discussed, while economists feel that they learn a great deal from such analyses (Shmeidler, 1989). We suggest that part of the knowledge generated by academic economists is case-based rather than rule-based. That is, instead of offering general rules or theories that should be contrasted with data, economists often analyze models that are "theoretical cases", which help understand economic problems by drawing analogies between the model and the problem. According to this view, economic models, empirical data, experimental results and other sources of knowledge are all on equal footing, that is, they all provide cases to which a given problem can be compared. In the paper is demonstrated DMSS engineering value driven approach within determination of the equilibrium points in the competitive trading modeled by the Edgworth box (Collopy & Hollingsworth, 2009; Ekeland, 1983).

The objective of the paper is to present an application of an innovative approach to construction of decision making support system. It combines two strategies – model-driven and decision-driven (Collopy & Hollingsworth, 2009). It is described a decision technology that could be used for realizing of suitable for PDA decision tool quantifies personnel judgment. It realizes evaluation (measurement) of human's, objectively oriented preferences as utility function.

Value and Utility Scales and Evaluations

One of the major problems faced by information and data-mining technologies is how to deal with uncertainty. The prime characteristic of Bayesian approach and methods is their explicit use of probability for quantifying uncertainty. Bayesian methods provide a practical method to make inferences from data using probability models for values we observe and about which we want to draw some hypotheses. The scientific point of view needs careful analysis of the terms *measurement*, formalization and admissible mathematical operations (Pfanzagl, 1971). This approach is closely related with utilization of theory of measurement and the utility theory. In the previous decades the Bayesian approach and the application of utility theory provoked serous critical debates about the so called normative or axiomatic approach in decision making. Let us remember the famous Allais paradox (Allais, 1953). Our position in regard to these fields of scientific areas is that we can use the stochastic programming and the Bayesian approach and utility theory from a prescriptive position by the use of the stochastic programming. This position is based on scientific and practical investigations made in Bulgarian Academy of Sciences in the previous three decades.

In complex processes and situations, there is a lack of measurements or even clearly identifiable scales for the basic heuristic information. Internal human expectations and heuristic are generally expressed by qualitative preferences. The common sources of information in such a basic level are the human preferences. According to social-cognitive theories, people's strategies are guided both by internal expectations about their own capabilities of getting results, and by external feedback (Bandura, 1986). Probability theory,

stochastic programming and expected utility theory address decision making under these conditions (Raiffa, 1968; Keeney & Raiffa, 1993).The mathematical description on such a fundamental level requires basic mathematical terms like sets, relations and operations over them, and their gradual elaboration to more complex and specific terms like value and utility functions, operators on mathematically structured sets as well, and equivalency of these descriptions. In this aspect of mathematical descriptions we enter the theory of measurements and scaling and utility theory**.**

Value Function

From practical point of view the empirical system of human preferences relations is a algebraic system with relations SR (***X*,(≈),(≻)**), where (≈) can be considered as the relation "indifferent or equivalent", and (≻) is the relation "prefer". We look for equivalency of the empirical system with the numbered system of relations SR (***R*-real numbers, (=), (>)**). The *"indifference"* relation (≈) is based on (≻) and is defined by $((x \approx y) \Leftrightarrow \neg((x \succ y) \vee (x \prec y)))$. Let **X** be the set of alternatives ($\mathbf{X} \subseteq \mathbf{R}^m$). A *Value function* is a function (u*: **X→R)** for which it is fulfilled (Keeney & Raiffa, 1993): $((x, y) \in \mathbf{X}^2, x \succ y) \Leftrightarrow (u^*(x) > u^*(y))$.

It is proved that for a finite set of alternatives and partial ordering there always exists such a function with precision up to monotonous transformation (Fishburn, 1970). In this manner we can move from the language of binary relations and preferences to the language of control criteria as objective value function. The assumption of existence of a value function *u*(.) leads to the "*negatively transitive*" and "*asymmetric*" relation (≻), "*weak order*". A "*strong order*" is a "*weak order*" for which is fulfilled $(\neg(x \approx y) \Rightarrow ((x \succ y) \vee (x \prec y)))$. The existence of a "*weak order*" (≻) over ***X*** leads to the existence of a "*strong order*" over ***X***/≈. Consequently the assumption of existence of a value function u(.) leads to the existence of: asymmetry $((x \prec y) \Rightarrow (\neg(x \succ y)))$, transitivity $((x \succ y) \wedge (y \succ z) \Rightarrow (x \succ z))$, axiom 5) and transitivity of the "indifference" relation (≈).

Utility Function and Measurement Scale

According to the *Utility theory* let ***X*** be the set of alternatives and ***P*** is a set of probability distributions over ***X*** and ***X*⊆*P***. A utility function *u*(.) will be any function for which the following is fulfilled:

$$(p \succ q\ ,\ (p,q) \in \mathbf{P}^2) \Leftrightarrow (\int u(.)dp > \int u(.)dq).$$

To every decision choice and act corresponds a discret probability distribution with finite domain of appearance of final alternatives (results). The notation (≻) expresses the preferences of DM over ***P*** including those over ***X*** (***X*⊆*P***). The interpretation is that the integral of the utility function *u*(.) is a measure concerning the comparison of the probability distributions *p* and *q* defined over ***X*** (figure 2).

There are different systems of mathematical axioms that give satisfactory conditions of a utility function existence. The most famous of them is the system of Von Neumann and Morgenstern's axioms (Fishburn, 1970):

(A.1) The *preferences* relations (≻) and (≈) are transitive, i.e. the binary preference relation (≻) is weak order**;**

(A.2) *Archimedean Axiom*: for all $p,q,r \in \mathbf{P}$ such that $(p \succ q \succ r)$, there is an $\alpha,\beta \in (0,1)$ such that $((\alpha\, p + (1\text{-}\alpha) r) \succ q)$ and $(q \succ (\beta p + (1\text{-}\beta) r))$;

(A.3) *Independence Axiom*: for all $p,q,r \in \boldsymbol{P}$ and any $\alpha \in (0, 1)$, then $(p \succ q)$ if and only if $((\alpha\, p + (1-\alpha)\, r) \succ (\alpha\, q + (1-\alpha)\, r))$.

Axioms (A1) and (A3) cannot give solution. Axioms (A1), (A2) and (A3) give solution in the interval scale (precision up to an affine transformation):

$$((p \succ q) \Leftrightarrow (\int v(x)\mathrm{d}p \succ \int v(x)\mathrm{d}q) \Leftrightarrow (v(x) = \boldsymbol{a}u(x) + \boldsymbol{b},\ \boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R},\ \boldsymbol{a} > 0,\ x \in \boldsymbol{X})).$$

The assumption of existence of a utility (value) function $u(.)$ leads to the "*negatively transitive*" and "*asymmetric*" relation ($\succ$) and to transitivity of the relation ($\approx$). So far we are in the preference scale, the *ordering scale*. The assumption of equivalence with precision up to affine transformation has not been included. In other words we have only a value function. For value, however, the mathematical expectation is unfeasible, but we underline that the mathematical expectation is included in the definition of the utility function. For this reason it is accepted that ($\boldsymbol{X} \subset \boldsymbol{P}$) and that $\boldsymbol{P}$ is a convex set:

$((q, p) \in \boldsymbol{P}^2 \Rightarrow (\alpha q + (1-\alpha)p) \in \boldsymbol{P}$, for $\forall \alpha \in [0,1])$.

Then utility $u(.)$ is determined in the interval scale (Fishburn, 1970):

Proposition 1. If $((x \in X \wedge p(x)=1) \Rightarrow p \in \boldsymbol{P})$ and $(((q, p) \in \boldsymbol{P}^2) \Rightarrow ((\alpha p + (1-\alpha)q) \in \boldsymbol{P}$, $\alpha \in [0,1]))$ are realized, then the utility function $u(.)$ is defined with precision up to an affine transformation: $(u_1(.) \approx u_2(.)) \Leftrightarrow (u_1(.) = a u_2(.) + b,\ a > 0)$.

Following from this proposition, the measurement of the preferences is in the *interval scale*. That is to say, this is a utility function. Now it is obvious why in practice the gambling approach is used to construct the utility function in the sense of von Neumann. The reason is that to be in the interval scale the set of the discrete probability distributions $\boldsymbol{P}$ have to be convex. The same holds true in respect of the set $\boldsymbol{X}$. The utility function is evaluated by the "gambling approach". This approach consists within the comparisons between lotteries. A "lottery" is called every discrete probability distribution over X. We denote as $\langle x, y, \alpha \rangle$ the simplest lottery: α is the probability of the appearance of the alternative x and $(1-\alpha)$ - the probability of the alternative y. In the practice, the utility measurement is based on the comparisons between lotteries as is shown in figure 3 (Raiffa, 1968; Keeney & Raiffa, 1993).

The weak points of the gambling approach are the violations of the transitivity of the preferences and the so called "certainty effect" and "probability distortion" identified by the Nobel prizeman Kahneman and Tversky. The violations of the transitivity of the relation equivalence ($\approx$) also lead to declinations in the utility assessment. All these difficulties explain the DM behavior observed in the Allais Paradox (Allais, 1953). Following the research of Kahneman and Tversky and the debates about the well known Allais paradox, extensions and further developments of von Neumann's theory were sought (Kahneman & Tversky, 1979). Among these theories the rank dependent utility (RDU) and its derivative cumulative Prospect theory are currently the most popular. In the RDU the decision weight of an outcome is not just the probability associated with this outcome. It is a function of both the probability and the rank the alternative. Based on empirical researches several authors have argued that the probability weighting function has an inverse S-shaped form, which starts on concave and then becomes convex.

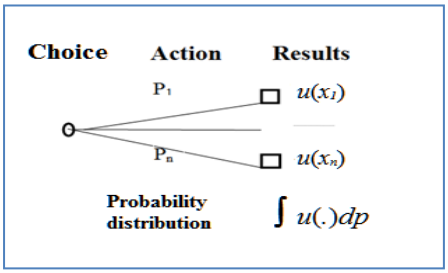


Figure 2.Acts, uncertainty and utility

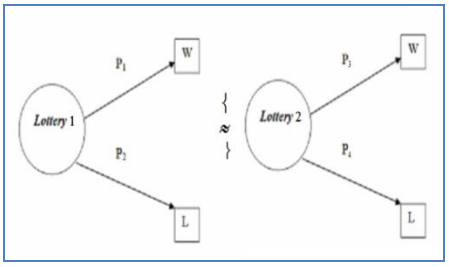


Figure 3. Gambling approach

Our approach to utility function evaluation includes the stochastic programming methods as prescriptive position in the decision making. We define two sets:

$\boldsymbol{A}_{u^*}=\{(\alpha,x,y,z)/(\alpha u^*(x)+(1-\alpha)u^*(y))>u^*(z)\}$,

$\boldsymbol{B}_{u^*}=\{(\alpha,x,y,z)/(\alpha u^*(x)+(1-\alpha)u^*(y))>u^*(z)\}$,

where $u^*(.)$ is DM's empirical utility. The analytical approximation of the utility function is constructed by recognition of the set $\boldsymbol{A}_u$ (Pavlov & Andreev, 2013). The proposed assessment is machine learning based on DM's preferences. The machine learning is a probabilistic pattern recognition ($\boldsymbol{A}_{u^*}\cap\boldsymbol{B}_{u^*}\neq\varnothing$) and the utility evaluation is a stochastic programming pattern recognition with noise (uncertainty) elimination**.** A experimental utility evaluation is shown in figure 4.

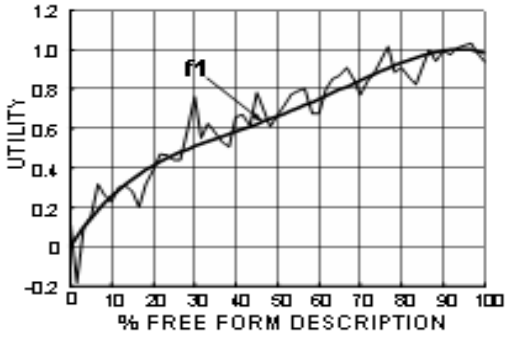


Figure 4. Stochastic utility evaluation

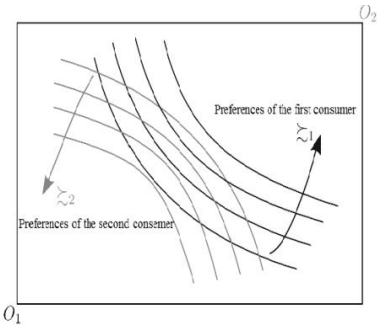


Figure 5. Consumer's demand's utility curves

The ciseaux line is probabilistic pattern recognition of the positive and negative preferences and smooth line is the utility function. The same approach is used for of value evaluation. The difference is only within the form of the sets $\mathbf{A}_{u^*}$ and $\mathbf{B}_{u^*}$**.** Let $\mathbf{A}_{u^*}$ and $\mathbf{B}_{u^*}$ be the sets:

$\mathbf{A}_{u^*}=\{(x,y)\in R^{2m}/\ (u^*(x))>u^*(y)\}$,

$\mathbf{B}_{u^*}=\{(x,\ y)\in R^{2m}/\ (u^*(x))<u^*(y)\}$.

If there is a function $\mathbf{F}(x,y)$ of the form $\mathbf{F}(x,y)=\mathbf{f}(x)-\mathbf{f}(y)$, positive over $\mathbf{A}_{u^*}$ and negative over $\mathbf{B}_{u^*}$, then the function $\mathbf{f}(x)$ is a value function, equivalent to the empirical value function $u^*(.)$. Such approach permits the use of stochastic "pattern recognition" for solving the problem. In the deterministic case it is true that $\mathbf{A}_{u^*}\cap\mathbf{B}_{u^*}=\varnothing$. In the probabilistic case it is

true that $\mathbf{A_{u^*}} \cap \mathbf{B_{u^*}} \neq \varnothing$ and here have to be used the probabilistic pattern recognition (Pavlov & Andreev, 2013).

The possibilities for practical preferences based stochastic evaluation of utility or value functions demonstrated in the paper permit a new position to value based decision making and elaboration of new information systems.

Value Driven Modeling: Edgeworth Box and Competitive Trade

The main pearpus of the paper is to demonstrate DMSS engineering value driven approach within determination of the equilibrium points in the competitive trading modeled by the Edgworth box. Competitive trade is a setting in which there are prices for two goods in question and many people who take these prices as given. A model for description the competitive trade is the Edgeworth Box. It merges the indifference map between the parties in the trade by inverting one of the agents diagram as is shown in figure 5. The demand functions or the utility functions which represent consumers' preferences are convex and continuous and are shown in figure 5. Given two consumers $\boldsymbol{O}_1$ and $\boldsymbol{O}_2$, two goods, and no production, all non-wasteful allocations can be drawn in the box shown in figure 6. Every point in the box represents a complete allocation of the two goods to the two consumers. Each of the two individuals maximizes his utility according to his preferences (Ekeland, 1983). The demand utility functions (figure 5) which represent consumers' preferences are convex and continuous, because in accordance with the theory the preferences in are continuous, monotone and convex (Ekeland, 1983).

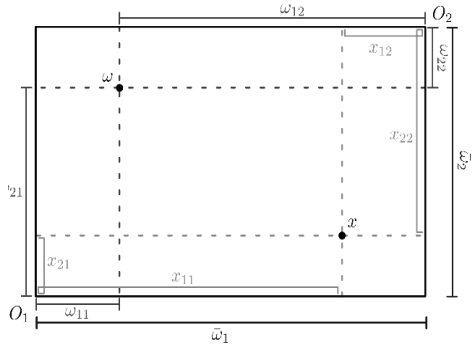


Figure 6. Edgeworth Box

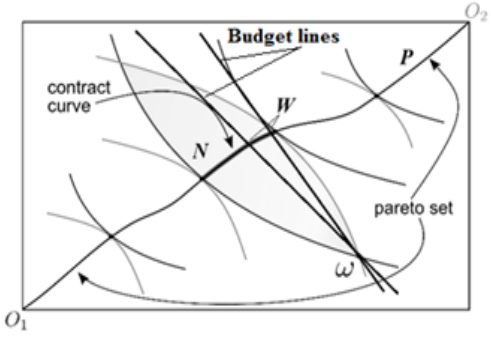


Figure 7**.** Contruct curves

Each consumer is characterized by an endowment vector, a consumption set, and regular and continuous preferences. The two consumers are each endowed (born with) a certain quantity of goods. They have locally non-satiated preferences and initial endowments:

$(\boldsymbol{w}_1, \boldsymbol{w}_2) = ((w_{11}, w_{21}), (w_{12}, w_{22}))$.

In the box the vector $w = (\overline{w}_1, \overline{w}_2)$ is the total quantities of the two goods:

$\overline{w}_1 = w_{11} + w_{12}$, $\overline{w}_2 = w_{21} + w_{22}$.

An allocation $\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2) = ((x_{11}, x_{21}), (x_{12}, x_{22}))$ represents the amounts of each good that are allocated to each consumer. A no wasteful allocation $\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is one for which is fulfilled:

$\overline{w}_1 = x_{11} + x_{12}$, $\overline{w}_2 = x_{21} + x_{22}$.

In terms of aggregate amounts of the two agents, the total amounts needs to be equal to the total endowment of the two goods. The consumers take prices of the two goods $\boldsymbol{p}$ =

(p_1, p_2) as given and maximize their utilities. The budget (income) set $B_i(\boldsymbol{p})$ of each consumer is given by: $\mathbf{B_i(p) = \{x_i \in R^2_+ \ / \ px_i \le pw_i\}, (i = 1, 2)}$, where ($\boldsymbol{px}_i$) and ($\boldsymbol{pw}_i$) mean scalar products. For every level of prices, consumers will face a different budget set. The locus of preferred allocations for every level of prices is the consumer's offer curve.

An allocation is said to be Pareto efficient, or Pareto optimal, if there is no other feasible allocation in the Edgworth economy for which both are at least as well off and one is strictly better off. The locus of points that are Pareto optimal given preferences and endowments is the Pareto set, noted as ***P*** in figure 7. The part of the Pareto set in which both consumers do at least as well as their initial endowments is the Contract curve shown in figure 7 and noted as ***N*** (kernel of market game).

We are interested in the equilibrium point(s) of the process of exchange where is fulfilled the Walrasian equilibrium (Ekeland, 1983). Walrasian equilibrium is a price vector ***p*** and an allocation ***x*** such that, for every consumer the prices (i.e. the terms of trade) are such that what one consumer (group of consumers) wants to buy is exactly equal to what the other consumer (group of consumers) wants to sell. In other words, consumers' demands are compatible with each other. We note the locus of points that are in Walrasian equilibrium as ***W*** (two points in figure 7). In still other words, the quantity each consumer wants to buy at the given market prices is equal to what is available on the market. The following inclusion is true in the Edgworth economy $\boldsymbol{P \supset N \supset W}$. In that sense a contract curve in the Edgeworth Box shows an exchange market in equilibrium and this is a particular representation of the Walrasian equilibrium theorem. The consumer's preferences are evaluated as value functions. In figure 8 are shown the indifference curves, the Pareto set ***P*** and the contract curve ***N***.

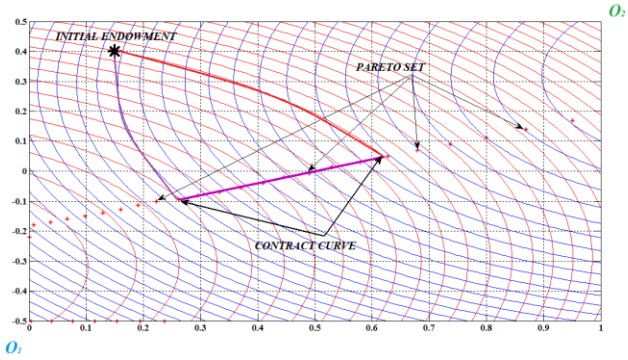


Figure 8. Edgeworth Box, Pareto set and contruct curves

The indifference curves in figure 8 are determined based on values functions evaluated by direct comparisons of couples of allocations $\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2) = ((x_{11}, x_{21}), (x_{12}, x_{22}))$. This is made through the discussed in the paper approach and algorithms for exact value function evaluation ($\mathbf{A_{u^*} \cap B_{u^*} = \varnothing}$). After that is maded quadratic approximation of the constructed value function. The divergence from the theoretical convex requirements is due to the finite number of learning points and to the uncertainty in the expressed consumer's preferences. In the experiment for determination of the set $\mathbf{A_{u^*}}$ and $\mathbf{B_{u^*}}$we used a finite number of preferences expressed for couples of allocations ($\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2)$, $\boldsymbol{y}=(\boldsymbol{y}_1, \boldsymbol{y}_2)$):

$\mathbf{A_{u^*} = \{(x,y) \in R^{2m} / \ (u^*(x)) > u^*(y)\}}$,

$\mathbf{B_{u^*} = \{(x, y) \in R^{2m} / \ (u^*(x)) < u^*(y)\}}$.

In that manner we can state and solve the market-clearing equilibrium in principle and we can determine the contract curve and the Walrasian set in the Edgeworth box. The set of

the Walrasian equilibriums $\boldsymbol{W}$ and the appropriate prices $\boldsymbol{p} = (p_1, p_2)$ are calculated based on the determined demand utility (value) functions and this is a meaningful prognosis of the market equilibrium. In that way can be forecast the competitive market equilibrium allocations $\boldsymbol{x}=(\boldsymbol{x}_1, \boldsymbol{x}_2) = ((x_{11}, x_{21}), (x_{12}, x_{22}))$ and the appropriate prices $\boldsymbol{p} = (p_1, p_2)$. The contract curves are specified on the individual consumers' preferences and show that there are possibilities to be made mutually advantageous trades. This means that one could unilaterally negotiate a better arrangement for everyone.

Conclusions

The analytical description of the expert's preferences as value or utility function will allow mathematically the inclusion of the decision maker in the value based model description of the complex system "Technologist-process. The indifference curves could be determined by utility or value function evaluation. The discussed previously in the paper stochastic procedures could be used for this purpose. In this case the learning points have to be defined as lotteries with Edgworth box allocations and consumers preferences in reference to learning triples of allocations.

The described methodology and procedures allow for the design of individually oriented information systems (Pavlov & Andreev, 2013). Our experience is that the human estimation contains uncertainty at the rate of [10, 30] %. Such systems allow for exact evaluation of the Pareto set $\boldsymbol{P}$, a reasonable determination of the contract curve $\boldsymbol{N}$ and calculation of the Walrasian set $\boldsymbol{W}$ and may be autonomous or parts of larger decision support system (Pavlov & Andreev, 2013; Ekeland, 1983). The demands functions could be evaluated by direct comparisons or by the gambling approach. In that manner the incomplete information is compensated with the participation of qualitative human estimations

References

1. A. Bandura, *Social foundations of thought and action: A social cognitive theory*, Englewood Cliffs, NJ: Prentice-Hall, 1986.
2. D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, *Econometrica*, *47*, 263-291, 1979.
3. D. Shmeidler, Subjective probability and expected utility without additivity, *Econometrica*, *57*(3), 571-587, 1989.
4. H. Raiffa, *Decision Analysis*, New York: Addison-Wesley Reading Mass, 1968.
5. I. Ekeland, *Elements d'économie mathématique*, Hermann, 1979, Russian translation: Mir, 1983.
6. J. Baron, Thinking and deciding (4th ed). Cambridge, MA: Cambridge University Press, 2008.
7. J. Pfanzagl, *Theory of Measurement*, Physical-Verlag, Wurzburg-Wien, 1971.
8. M. Allais, *Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine*, *Econometrica*, *21*, 503-546, 1953.
9. P. Collopy, P. Hollingsworth, Value-driven design. *AIAA Paper* 2009-7099, *American Institute of Aeronautics and Astronautics*, Reston, VA, 2009.
10. P. Fishburn, *Utility theory for decision-making*, New York, Wiley, 1970.
11. R. Keeney, H. Raiffa, *Decision with multiple objectives: Preferences and value trade-offs*, Cambridge & New York: Cambridge University Press, 1993.
12. Y. Pavlov and R. Andreev, *Decision control, management, and support in adaptive and complex systems: Quantitative models*, Hershey, PA: IGI Global, 2013.

Modelling and Prediction of the Electricity Consumption

Andrey Bachvarov[1], Petko Ruskov[2], Kaloyan Haralampiev[2]

[1]Balkans Investment Consulting Agency Ltd., andrey.bachvarov@bica-bg.com
[2]Sofia University, petkor@fmi.uni-sofia.bg; k_haralampiev@hotmail.com

Abstract. The goal of the paper is to model, analyze and predict the electricity consumption of the electric utility company at west Bulgaria. To achieve this goal this paper combines the results of academic research of applicable best practices, empirical tests based on real data obtained with the kind permission of a local distribution company and the analytical results obtained through the use of advanced statistical software. Our research greatly relies on data provided by the Operations department of "CEZ Electro Bulgaria" AD and subsequent sophisticated data mining analysis in order to detect and expose the consumption. The results include electricity consumption models developed through the use of the CRISP-DM research methodology using IBM SPSS Modeller predictive models and tools. The generated models, and predicted results can be readily applied by utility companies, their partners and consultants for future energy efficiency processes innovation and business transformation.

Keywords. Electricity consumption, predictive model analysis.

1. Introduction

Electricity consumption in Bulgaria and in the world is in high demand and at a premium cost to the consumer. Having a satisfactory energy supply is vital for the community. The electricity consumption prediction is an essential component of the optimization tools adopted by energy companies for electricity power system scheduling. A small improvement in load forecasting can bring substantial benefits by reducing production costs as well as increasing trading advantages, especially during peak periods. The electric utility companies are accountable for maintaining and improving their service continuously. To facilitate better planning, utility companies maintain databases that capture the energy consumption and usage patterns of applications. These databases are then used to identify the trend and the usage of domestic energy [5, 9].

As the recession stays in the air over us, energy consumers are becoming smarter and more sensible of their electricity consumption [1, 15, 16]. A prediction model for the electricity consumption at a local level is of great interest to today's producers, distributors and consumers [4]. A model needs to be created so that it can predict electricity consumption and prices on a local level, implying a single arrangement.

From the other side the innovation is a driving force of prosperity and contributes to increasing the production and the standard of living [10, 12]. It is considered to be a critical component of business productivity and competitive survival [11]. Technological innovations present vast opportunities for product innovation which is introduction of new types of goods and services for the external market, process innovation which is enhancement of internal production processes for goods and services.

The objective of this paper is to present statistical prediction models and to evaluate the impact of electricity consumptions of the electric utility company at west Bulgaria. In it

the method used for grid planning is based on statistical analysis of the existing data for each local area the maximum load is projected to change proportional to changes in the aggregated national electricity consumption. Based on a, IBM SPSS modeler valuable models are developed. The findings also revealed that the regression model reflected the total opposite for dollar amounts over time. The combination of kilowatt hour and time in one model is more significant than them apart. The developed models are utilized to evaluate the impact of the past consumptions and temperature to predict the future electricity consumption. Long term projections of daily or hourly electricity consumption in local areas are important for planning of the transmission grid [2].

2. Modeling procedure in IBM SPSS Modeler

The paper combines the results of academic research of applicable best practices, empirical tests based on real data obtained with the kind permission of a local distribution company and the analytical results obtained through the use of advanced statistical software [3, 14]. Proposed model for research and forecasting electricity consumption depending on past consumption data and temperature are based on processing large amounts of data provided by CEZ Electro Bulgaria AD and the average temperature in the respective localities [3]. The IBM SPSS Modeler was used for processing the raw data for 433,558 subscribers of CEZ and 279 variables and the results are presented [6, 7].

The raw data consist of 279 variables for 433 558 consumers of CEZ. For each user we have the reporting date and meter readings, separately for day and night tariff, from May 2010 to February 2013.

Based on the reporting dates we calculate the length of the period in days (number of days for given consumption for the given month). Based on the meter readings we calculate consumed energy. Then we calculate the average daily consumption (for 1 day of the given month). These calculations are made because the metering of a user consumption is not exactly for the entire month (from first to the last date of the month), and it is toward a specific date, different for different users. Thus we add new 198 variables consist of average day and night consumptions to the database.

We calculate the total amount of the average daily consumption of all users for each month separately. For evaluation of the total consumption for a given month we multiply the January average amount by 31, the February average amount by 28(29), etc. Thus we obtain two time series, for total daily and total nighttime consumption, respectively, for the period from May 2010 to February 2013.

Then we model these two time series with IBM SPSS Modeler and make forecasts to the end of 2013.The most significant consideration is that these quantitative models had to be built on the basis of information available to the organization. This is information that is collected on a daily basis in corporate databases as a result of current activities.

Definition of Corporate Sustainability

For time series analysis and making forecasts IBM Modeler uses the Time Series node which consists of:

- Univariate exponential smoothing (univariate modelling uses only one time series, i.e. a separate model is made for each time series.),
- ARIMA (Autoregressive Integrated Moving Average),
- Transfer function models.

The procedure includes an Expert Modeler that identifies and estimates an appropriate model for each dependent variable series. [8, p. 345].

Users can let the Expert Modeler select a model for them from:

- All models (default).
- Exponential smoothing models only.
- ARIMA models only [8, p. 359].

All Models Expert Model

In this case, the Exponential Smoothing and ARIMA expert models are computed, and the model with the smaller normalized BIC is chosen. [8, p. 360]

BIC means Bayesian Information Criterion and it is calculated by the following equation:

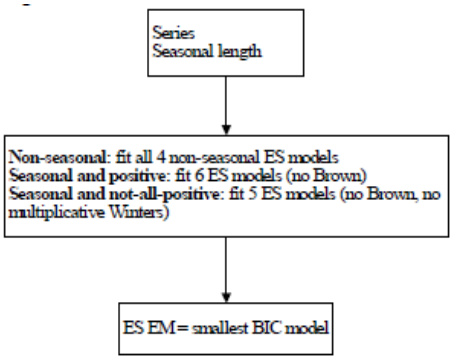
$$\text{Normalized } BIC = \ln(MSE) + k\frac{\ln(n)}{n}$$

where MSE is Mean Squared Error:

$$MSE = \frac{\Sigma\left(Y(t) - \hat{Y}(t)\right)^2}{n-k}$$

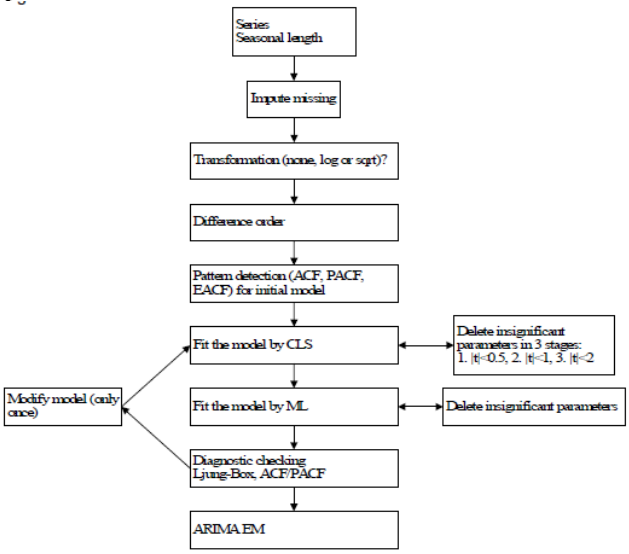
Y(t) is the original time series, $\hat{Y}(t)$ is the predicted time series, k is number of parameters in the model, and n is number of non-missing observations.

Exponential Smoothing Expert Model



Source: IBM SPSS Modeler 15 Algorithms Guide, p. 359

ARIMA Expert Model



Source: IBM SPSS Modeler 15 Algorithms Guide, p. 360

As could be seen it remains only to choose the actual time series and let the IBM SPSS Modeler to do the rest of the work:

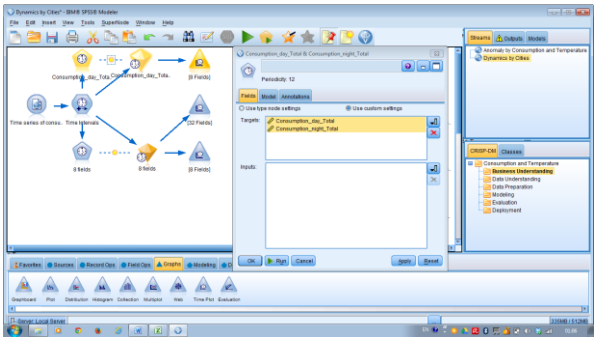


Figure 1. Model of the dynamics (dynamics of total consumption is marked in yellow)

As a result of modeling of the dynamics of the total daily and total nighttime consumption we obtain that for both time series the best model is Simple seasonal:

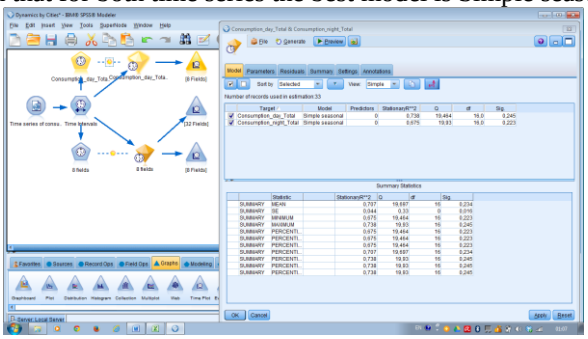


Figure 2. The best models of the dynamics of the total daily and total nighttime consumption

The best models for all provided consumption data together with the forecasts for the next 10 months to the end of 2013 and the confidence intervals of the models and forecasts are presented graphically in Figure 3:

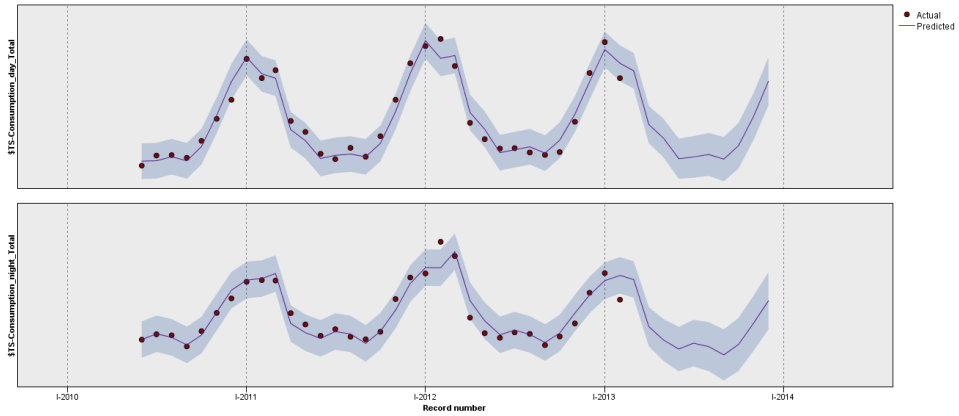


Figure 3. Models and forecasts of the total daily and total nighttime consumption for all customers of CEZ, for which we have data

Besides graphically, models, forecasts and confidence intervals can be presented in tables, upon request by the utility company.

Because for a part of users (about half) we have also data for the postal codes, we can trace the dynamics of consumption by settlements for those consumers who have a postal code. In the database contains data for four district centers in Western Bulgaria – Sofia, Pernik, Vratsa and Lovech.

As a result of modeling of the dynamics of the daytime and nighttime consumption by cities we obtain that for the eight time series the best model is again Simple seasonal.

The best models for daily and nighttime monthly consumption of these cities together with the forecasts for the next 10 months to the end of 2013 and the confidence intervals of the models and forecasts are presented graphically in Figure 4:

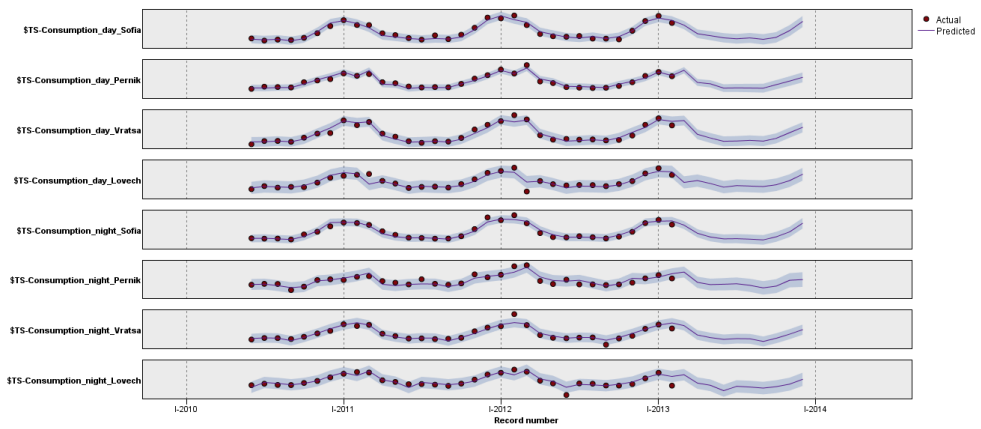


Figure 4. Models and forecasts of the daytime and nighttime consumption for customers of CEZ of the four cities

The Figure 4 shows that the peak consumption of the households is in the winter and the consumption is the least during the summer. The use of air conditioners during the summer months, which is a trend for business users, is not yet common practice in everyday life.

Henceforth, the analysis continues with more emphasis on the capital city of Sofia, but likewise it could be done for the other three cities for which the data are available.

First, from Figure 5 we can subtract only the models and forecasts for Sofia – Figure 5:

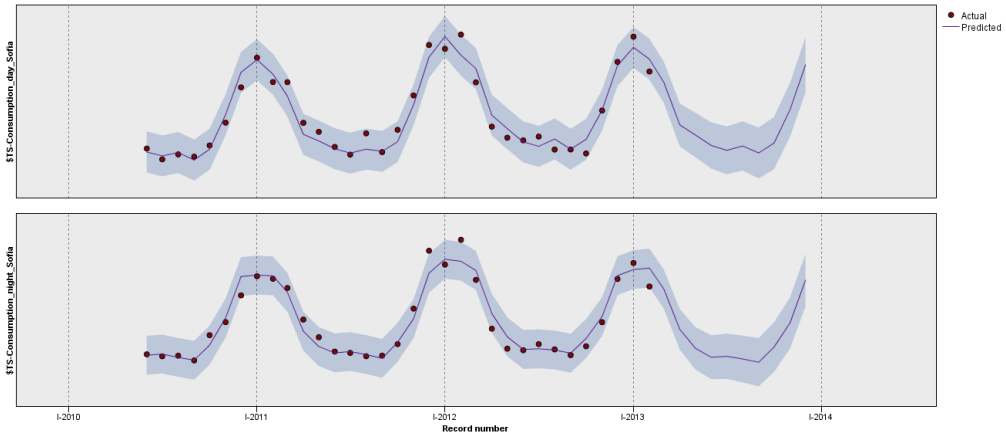


Figure 5. Models and forecasts of daytime and nighttime consumption for customers of CEZ from Sofia

Second, we took data from the National Institute of Meteorology and Hydrology of the Bulgarian Academy of Sciences for the average daily and average nighttime temperatures in the nine district centers in Western Bulgaria where CEZ operates, plus Botevgrad, for the period from January 2011 to August 2012. This allows us for that short time period to include the average daily and nighttime temperatures for the respective months as predictors in the model shown in Figure 6:

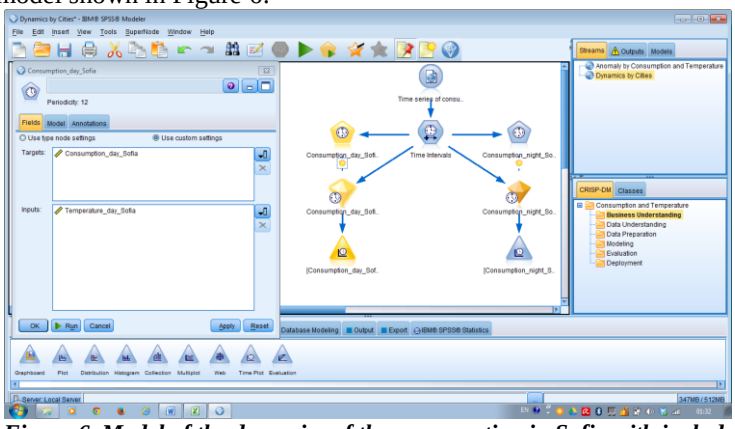


Figure 6. Model of the dynamics of the consumption in Sofia with included temperatures (dynamics of daily consumption is marked in yellow)

As a result of modeling of the dynamics of the daily consumption in Sofia with included average daily temperatures we obtain that the best model is ARIMA (0,1,0) – Figure 7:

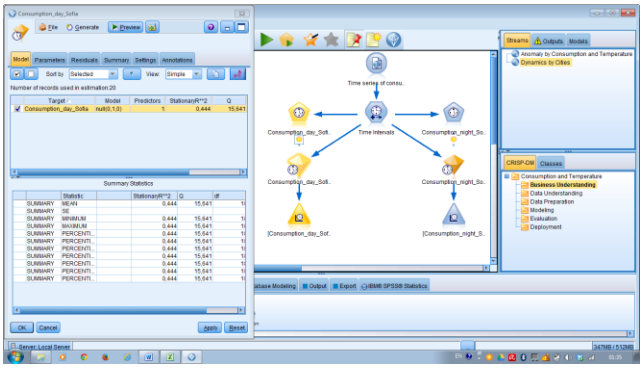


Figure 7. The best model of the dynamics of daily consumption in Sofia with included average daily temperatures

When we include the temperatures as predictors in the model we are unable to make forecasts (because there are no predictions for the temperatures themselves) so we present graphically only the best model with confidence intervals of the model – Figure 8:

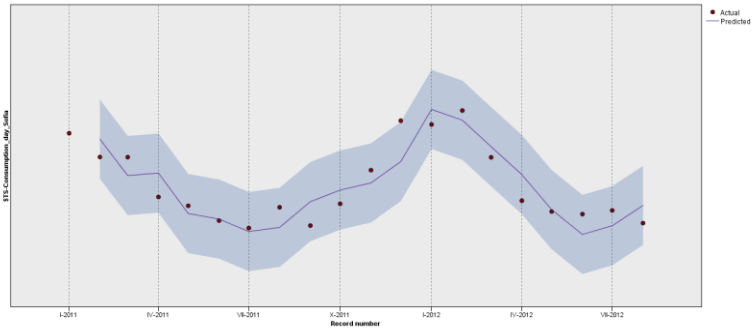


Figure 8. Model of the daily consumption of customers of CEZ from Sofia with included average daily temperatures

In the same manner we have modeled the nighttime consumption in Sofia:

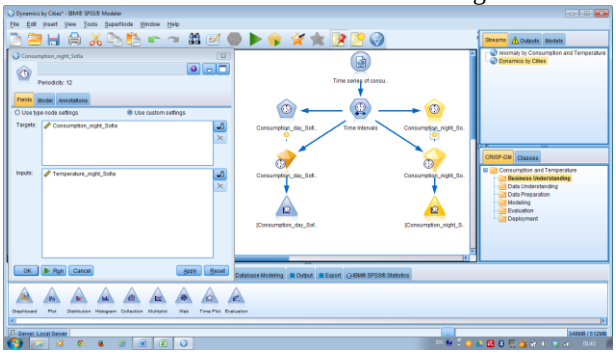


Figure 9. Model of the dynamics of consumption in Sofia with included temperatures (dynamics of nighttime consumption is marked in yellow)

As a result of modeling of the dynamics of the nighttime consumption in Sofia with included average nighttime temperatures we obtain that the best model is Simple seasonal:

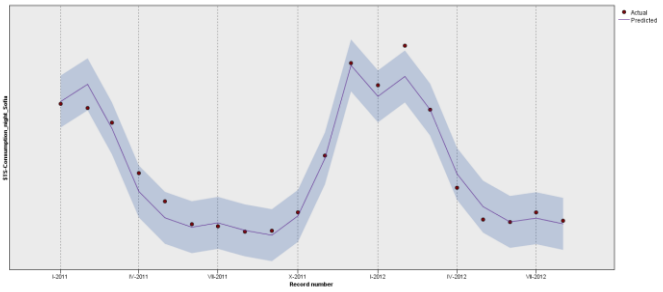


Figure 10. Model of nighttime consumption of the customers of CEZ from Sofia with included average nighttime temperatures

For comparison of the models of the dynamics of the daytime and nighttime consumption in Sofia, with and without incorporated temperatures, the results are combined in Figures 11:

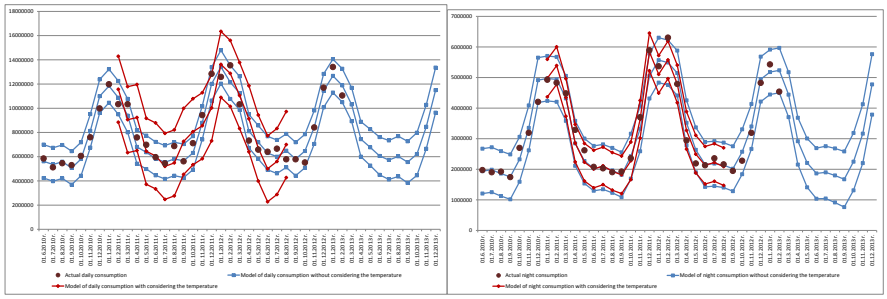


Figure 11. Comparison of the models of the dynamics of daily (left) and night time consumption in Sofia with and without incorporated temperatures

When we model the dynamics of daily consumption in Sofia better results are obtained without taking into account the average daily temperatures, while modeling the dynamics of nighttime consumption results with and without taking into account the average nighttime temperatures are practically identical. This is an expected result because R-square of the two models of nighttime consumption in Sofia (with and without the average monthly temperatures) is practically the same – 0.763 and 0.756 respectively, while R-square of the model of the daily consumption in Sofia including temperatures is almost two times less than R-square of the model of daily consumption in Sofia which not includes temperatures – 0.444 and 0.803 respectively.

Assuming that the user consumption profile (demographic and social) for categories of customers is approximately identical for all local areas we estimate the daily consumption in local areas. Using it and assuming that profiles per category are unwavering in the future, a baseline projection of local consumption is generated. Applying the latest forecast of the electricity consumption at the regional level, consumption in local areas develops quite differently.

Using the estimated data and the consumption profiles for categories of electricity customers in the aggregated area, the model also calculates the maximum and minimum daily consumption in local areas. For grid planning a high consumption in several consecutive days is more important than extreme consumption in individual days, and to

evaluate consumption in consecutive days a modeling of the entire consumption profile is required. Assuming that the modeling inaccuracy is constant over time, changes in the maximum and minimum consumption are calculated. In overall, consumption in peak days is expected to increase more than consumption in average and minimum days and these contrasts between local areas and for some areas the opposite is the case.

The model takes into account a local data of categories of customers, only. Using the models, the electricity consumptions are assumed to be constant and the profiles for categories of customers are assumed to be equal to the profiles for the aggregated area and constant. Profiles for categories of customers may change over time, profiles for categories in local areas may differ from the profiles for the aggregated area, and the data of categories may change. That is, the model gives a baseline projection, only, and should be supplemented with specific local climate, social and demographic information, e.g. on the development of a new dwelling area, increased consumption of a specific customer, or changes in the consumption profile for categories of customers. Compared to the methodology described in the literature our model has the advantages of more inside data and iterations describing the entire consumption duration curve, allowing different changes in average consumption in different local regions, and changes in the shape of the local load duration curve.

3. Conclusion

Traditional business practice will not fill the energy system gap. Only the kind of business analytics, systemic, transformative changes to the energy system and following innovation of processes and business models will help the society to achieve a sustainable and secure energy future.

This paper presented Time Series Algorithms models to predict residential building energy consumption. Data and results of the models show that consumption profiles for different categories of customers contribute differently to the aggregated consumption profile and the relative weight of categories differs between local areas, aggregated consumption profiles for local areas are quite different. The results include electricity consumption models developed through the use of the CRISP-DM research methodology using IBM SPSS Modeller predictive models and tools. The generated models, and predicted results can be readily applied by utility companies, their partners and consultants for future energy efficiency processes innovation and business transformation.

From measurements of local consumption we know that:

• Consumption profiles differ between local areas;

• Consumption by categories of customers contributes differently to the aggregated consumption profile;

• The weight of categories of customers differs between local areas.

That is, specific local conditions are not considered. Specifically, the study serves to highlight a number of potentially significant issues for future work:

• Knowing uncertainties within the models require further field studies of users characteristics, indoor temperatures, and consumptions behavior;

• There is need for further modeling effort to investigate the uncertainties to the most influential factors used to describe the day and hour consumptions.

Acknowledgement

The authors are grateful to the Marketing & Sales support department of "CEZ Electro Bulgaria" AD for the explicit sample data base and helpful information.

The work on this paper has been sponsored by IBS Bulgaria, IBM Premier Business partner

4. References

1. Amit R., Christoph Zott, Creating Value Through Business Model Innovation, MIT SLOAN MANAGEMENT REVIEW, Spring 2012, Reprint 53310, Vol. 53, No 3, pages 41-49.
2. Bachvarov A., Petko Ruskov, Kaloyan Haralampiev, Electricity end users' consumption patterns in different west Bulgarian locations, UNITECH'12 International scientific conference, 16 – 17 November 2012, Gabrovo.
3. CEZ, CEZ Electro Bulgaria AD, http://www.cez.bg/en/home.html, http://www.cez.bg/en/customer-service/ways-of-payment.html.
4. Martin Greene, P.E., and Jin-Lee Kim, Prediction Model of California Residential Buildings Energy Consumption, ICSDEC 2012, © ASCE 2013.
5. Carl Haigney, Customer Relationships and the Economy, Shaping a New Era in Energy, Accessed on 24 January 2014, www.UtilitiesProject.com.
6. IBM Corp., CRISP-DM 1.0, Step-by-step data mining guide, © Copyright IBM Corporation 2010.
7. IBM Corp., Switching perspectives, creating new business models for a changing world of energy, 2010.
8. IBM SPSS Modeler 15 Algorithms Guide , Copyright IBM Corporation 1994, 2012.
9. M. Kavgic, D. Mumovic, A. Summerfield, Z. Stevanovic, O. Ecim-Djuric, Uncertainty and modeling energy consumption: Sensitivity analysis for a city-scale domestic energy model, Energy and Buildings 60 (2013) 1–11.
10. Kiron D., Pamela Kirk Prentice, Renee Boucher Ferguson, Innovating With Analytics, MIT SLOAN MANAGEMENT REVIEW, VOL.54 NO. 1, 2012.
11. LaValle S., Michael Hopkins, Eric Lesser, Rebecca Shockley and Nina Kruschwitz, Analytics: The new path to value, How the smartest organizations are embedding analytics to transform insights into action, IBM Institute for Business Value, MIT Sloan Management Review, 2010.
12. Richard. K. Lester and David M. Hart Unlocking Energy Innovation: How America Can Build a Low-Cost, Low-Carbon Energy System, MIT Press, November 2011.
13. Morelli T., Shearer C., Buecker A., IBM SPSS predictive analytics: Optimizing decisions at the point of impact, IBM redpaper 4710, Copyright IBM Corp. 2010.
14. Ruskov P., A. Bachvarov, Disruptive innovations for the energy system, Anniversary Scientific Conference with International Participation "40 Years Department of Industrial Automation", 18 March 2011, Sofia, Bulgaria, pp. 141-144.
15. Lukas G. Swan, V. Ismet Ugursal, Modeling of end-use energy consumption in the residential sector: A review of modeling techniques, Renewable and Sustainable Energy Reviews 13 (2009) 1819–1835, 2009.
16. Marc Teerlink and Michael Haydock, Customer analytics pay off, Driving top-line growth by bringing science to the art of marketing, Copyright IBM Corporation 2011.

Comparison of Commercial CFD Software Packages

Dragan Pavlovic[1], Milena Todorovic[1], Milos Jovanovic[1], Pedja Milosavljevic[1]

[1] University of Nis, Faculty of Mechanical Engineering, 18000 Nis, Serbia
draganpavlovic10369@gmail.com

Abstract. This paper presents the possibilities of Computational Fluid Dynamics (CFD) application, its advantages, modelling process and activities, as well as the review of the CFD Software packages. CFD is a branch of fluid mechanics that uses numerical methods and algorithms to solve and analyze problems that involve fluid flows and related phenomena (heat transfer, mass transfer, chemical reactions) by solving numerically the respective governing equations. The results of CFD analysis are relevant engineering data that are used in conceptual studies of new designs, detailed product development and redesign. The paper also summarizes the features and application of the CFD Software packages like Phoenics, Fluent and CFX and discusses their structure features, special modules, mathematical models and successful application areas. At the end the paper it points out the possible research direction for CFD in the future focusing on the development of mathematical models, project transformation, new equipment and their matching application with technological software.

Keywords. Computational Fluid Dynamics, Phoenics, Fluent, CFX, modelling and simulation techniques.

1. Introduction

Computational fluid dynamics, usually abbreviated as CFD, is a branch of fluid mechanics that uses numerical methods and algorithms to solve and analyze problems that involve fluid flows. Nowadays the computers are used to perform required calculations to simulate the interaction of liquids and gases with surfaces, defined by boundary conditions. Better solutions can be achieved by using high-speed supercomputers. Current development and research results lead to improvement of accuracy and speed of complex simulation scenarios such as transonic or turbulent flows.

Basically, almost all CFD problem analysis lies on the basis of the Navier–Stokes equations, which define any single-phase (gas or liquid, but not both) fluid flow. In order to yield the Euler equations Navier–Stokes equations can be simplified by removing terms that describe viscous actions. Further simplification, by removing terms describing vorticity, brings the full potential of equations. Finally, for small perturbations in subsonic and supersonic flows (not transonic or hypersonic) these equations can be linearized to yield the linearized potential equations.

Historically, methods were first developed to solve the Linearized potential equations. Two-dimensional (2D) methods, which use conformal transformations of the flow around cylinder to the flow around an airfoil were developed in the 1930s [1]. The computer power available paced development of three-dimensional methods. The first work using computers to model fluid flow, as governed by the Navier-Stokes equations, was performed at Los Alamos National Labs, in the T3 group [2], [3].This group was led by Francis H. Harlow, who is widely considered as one of the pioneers of CFD. From 1957 to late 1960s, this group

developed a variety of numerical methods to simulate transient two-dimensional fluid flows, such as Particle-in-cell method [4],Fluid-in-cell method [5], Vorticity stream function method [6], and Marker-and-cell method [7]. Fromm's vorticity-stream-function method for 2D, transient, incompressible flow was the first treatment of strongly contorting incompressible flows in the world.

The first paper with three-dimensional model was published by John Hess and A.M.O. Smith of Douglas Aircraft in 1967[8]. This method discretized the surface of the geometry with panels, giving rise to this class of programs being called Panel Methods. Their method itself was simplified, in that it did not include lifting flows and hence was mainly applied to ship hulls and aircraft fuselages. The first lifting Panel Code (A230) was described in a paper written by Paul Rubbert and Gary Saaris of Boeing Aircraft in 1968[9].

2. CFD Application and its Advantages

CFD is the system analysis that involves fluid flow, heat transfer and associated phenomena such as chemical reactions by means of computer-based simulation. The technique is very useful and includes a wide range of industrial and non-industrial applications areas. Some examples are: chemical process engineering: mixing and separation, multiphase systems, aerodynamics of aircraft and vehicles, hydrodynamics of ships, power plants, turbo machinery, electrical and electronic engineering, external and internal environment of buildings, marine engineering, environmental engineering, hydrology and oceanography, meteorology, and biomedical engineering (blood flows through arteries and veins). CFD is becoming a vital component in the design of industrial products and processes [10].

There are several unique advantages of CFD over experimental-based approaches to fluid systems design: substantial reduction of lead times and costs of new designs, ability to study where controlled experiments are difficult or impossible to perform, ability to study systems under hazardous conditions at and beyond their normal performance limits, and practically unlimited level of detail of results. The variable cost of an experiment, in terms of facility hire and/or man-hour costs, is proportional to the number of data points and the number of configurations tested. In contrast CFD codes can produce extremely large volumes of results at virtually no added expense and it is very cheap to perform parametric studies, for instance to optimize equipment performance [10]. One of the advantages in using CFD Software is also the speed of the simulation. CFD simulations can be executed in a short period of time. Quick turnaround means engineering data can be introduced early in the design process. Also, what is very important is ability to simulate real and ideal conditions. Many flow and heat transfer processes cannot be easily tested - e.g. hypersonic flow at Mach 20, nuclear accident etc.; CDF provides the ability to theoretically simulate any physical condition. Among other, CFD allows great control over the physical process and provides the ability to isolate specific phenomena for study. For example, a heat transfer process can be idealized with adiabatic, constant heat flux, or constant temperature boundaries. One of the very important properties of CFD is that it can give comprehensive information. Experiments only permit data to be extracted at a limited number of locations in the system (e.g. pressure and temperature probes, heat flux gauges, LDV, etc.).CFD allows the analyst to examine a large number of locations in the region of interest, and yields a comprehensive set of flow parameters for examination [11].

But there are some limitations in using CFD modelling. Regarding physical models, CFD solutions rely upon physical models of real world processes (e.g. turbulence, compressibility, chemistry, multiphase flow, etc.). The solution that are obtained through

CFD can only be as accurate as the physical models on which they are based. There are also limitations regarding physical boundary condition: as with physical models, the accuracy of the CFD solution is only as good as the initial/boundary conditions provided to the numerical model. Regarding numerical errors: solving equations on a computer invariably introduces some numerical errors such as:

- Round-off error - represents errors due to finite word size available on the computer;

- Truncation error - represents error due to approximations in the discretization schemes.

Round-off errors will always exist (though they should be small in most cases) and truncation errors will go to zero as the grid is refined. In that case mesh refinement is one way to deal with truncation error and/or schemes of greater accuracy are used [11].

3. Selecting CFD Software – Modelling Process Features

There are many commercial CFD softwares used in engineering, such as PHOENICS (it is the first commercial CFD software), STAR-CD, ANSYS FLUENT/CFX and so on. All CFD softwares have three main structures which are Pre-Processer, Solver and Post-Processor [12].

No matter what kind of CFD software is, the main processes of simulation are the same. Setting up governing equations is the precondition of CFD modelling; mass, momentum and energy conservation equation are the three basis governing equations. After that, Boundary conditions are decided as different flow conditions and a mesh is created. The purpose of meshing model is discretized equations and boundary conditions into a single grid. A cell is the basic element in structured and unstructured grid. The basic elements of two-dimensional unstructured grid are triangular and quadrilateral cell. Meanwhile, the rectangular cell is commonly used in structured grid. In 3D simulation, tetrahedral and pentahedral cells are commonly used in unstructured grid and hexahedral cell is used in structured grids. The mesh quality is a precondition for obtaining the reasonably physical solutions and also represents the ability of the simulation. The more nodes resident in the mesh, the greater the computational time to solve the aerodynamic problem concerned, therefore creating an efficient mesh is indispensable. Numerical methods that are used to discretize equations are: Finite Different Method (FDM), Finite Element Method (FEM) and Finite Volume Method (FVM). FVM is widely used in CFD software such as Fluent, CFX, PHOENICS and STAR-CD, to name just a few. Compared with FDM, the advantages of the FVM and FEM are that they are easily formulated to allow for unstructured meshes and have a great flexibility so that can apply to a variety of geometries [13]. Desired features in the modelling process in different CFD packages are divided in following groups according the modelling process itself [14].

Pre-processor. Pre-processor consists of the input of a flow problem to a CFD program by means of an operator-friendly interface and the subsequent transformation of this input into a form suitable for use by the solver. The user activities at the pre-processing stage involve: definition of the geometry of the region of interest: the computational domain, grid generation-the sub-division of the domain into a number of smaller, non-overlapping sub-domains: a grid (or mesh) of cells (or control volumes or elements), selection of the physical and chemical phenomena that need to be modelled, definition of fluid properties, specification of appropriate boundary conditions at cells which coincide with or touch the domain boundary. The solution to a flow problem (velocity, pressure, temperature, etc.) is defined at nodes inside each cell. The accuracy of a CFD solution is governed by the number

of cells in the grid. In general, the larger the number of cells the better the solution accuracy. Both the accuracy of a solution and its cost in terms of necessary computer hardware and calculation time are dependent on the fineness of the grid. Optimal meshes are often non-uniform: finer in areas where large variations occur from point to point and coarser in areas with relatively little change. Over 50% of the time spent in industry on a CFD project is devoted to the definition of the domain geometry and grid generation [15-18].

Solver. There are three distinct streams of commercial solution techniques: finite difference, finite element and spectral methods. In outline the numerical methods that form the basis of the solver perform the following steps: approximation of the unknown flow variables by means of simple functions, discretisation by substitution of the approximations into the governing flow equations and subsequent mathematical manipulations, solution of the algebraic equations.

Post-processor. As in pre-processing a huge amount of development work has recently taken place in the post-processing field. Owing to the increased popularity of engineering workstations, many of which have outstanding graphics capabilities, the leading CFD packages are now equipped with versatile data visualization tools. These include: domain geometry and grid display, vector plots, line and shaded contour plots, 2D and 3D surface plots, particle tracking, view manipulation (translation, rotation, scaling etc.), color postscript output. More recently these facilities may also include animation for dynamic results display and in addition to graphics all codes produce trustily alphanumeric output and have data export facilities for further manipulation external to the code.

One of the desired features is also the user-friendliness. The capability that the user is able to operate the software is determined by the design of the user interface. The more friendly a package appears, the shorter learning curve will be and the quicker users will be able to obtain meaningful results. The Graphical User Interface (GUI) provides a very easy introduction to the software and enables basic problems to be set up very quickly using both mouse and keyboard. Less problems can be defined using a command language which "sits beneath" the GUI. This enables files that are normally produced automatically using the top level menu, to be written (and edited) by the user using a very high level and logical language. The command language offers users much more freedom in defining geometries, meshes, boundary conditions, etc. However, if the command language is still not flexible enough, it is necessary for the user to access User Fortran Routines. These are sub-sections of the main CFD code in which users can write Fortran routines to run their simulation. For example, to define new physical models, specify complex boundary conditions, define additional output variables, etc.

User-support is also defined as feature that is crucial for selecting the CFD packages. When using CFD software to address difficult problems, it is helpful if good user support is available via an immediate, rapid response method, such as fax, email or telephone. Some vendors also set up email user groups which enable users to send an advice request to all current users of that CFD software packages. Most companies offer training courses to new and potential users enabling them to overcome some of the initial difficulties and get the most from the software in a shorter period of time. To gain an insight into the quality of user-support, there is much to be gained from visiting potential suppliers before selecting a software vendor [15-18].

4. Review of the CFD Software Packages

Several packages were chosen for investigation. These were considered according a list of desired features that are presented above (Table I and II). Later further analysis

considered the most widely used codes such as: CFX, FLUENT and PHOENICS in more detail (Table III), as well as comparison of the process features between top selling vendors Ansys CFD and SolidWorks (Table IV).

Table 1. The names and the data on the chosen CFD packages that are taken into consideration in this comparative review [19-27]

Company	Code	Operating system	Hardware requirements
ANSYS, Inc. www.ansys.com	CFX 14.5	Windows, Linux, Solaris	Windows XP SP3 (min 2GB RAM) 100GB hard disk free space, 2GB additional hard disk space needed Windows 7 (min 4GB RAM) 5GB hard disk free space, 10GB additional hard disk space needed
	FLUENT 13		Windows XP SP3 (min 512MB RAM) 100GB hard disk free space Windows 7 (min 2GB RAM) 100GB hard disk free space
CHAM Ltd. www.cham.co.uk	Phoenics	Windows, Linux, Solaris	Windows XP (min 512MB RAM), Windows Vista/7 (min 1GB RAM) 1.2-2.2GB hard disk free space
Flow Science, Inc. www.flow3d.com	FLOW-3D/MP	Windows, Linux	both Workstations and clusters, Redhat Enterprise Linux (5 or 6) or SUSE Linux 11 large shared NFS disk
CD adapco Group www.cd-adapco.com	STAR-CCM+	Windows, Linux, Solaris	Cluster Platform: Windows or Linux
	STAR-Cast		
	STAR-CD 4.06		
SolidWorks Corp. www.solidworks.com	SolidWorks Simulation 2013	Windows, Linux	Windows 7 (32- or 64-bit), Windows 8 (64-bit) or Windows Vista 2GB RAM (min) 5GB disk space free (min)

Table 2. Comparison of the CFD packages according to desired features [19-27]

Feature	Cfx	Fluent	Phoenics	Star-sd
Discretisation technique	FEM	FVM	FVM	FVM
Mesh type	STR	STR	STR	UNS
2D and 3D	Y	Y	Y	Y
Friendly interface	Y	Y	Y	Y
Body-fitted coordinates	Y	Y	Y	N/Y
Turbulence models: k - ε	Y	Y	Y	Y
Multi-phase/species	Y	Y	Y	Y
Access to Fortran user-routines	Y	Y	Y	Y

FEM - finite element method, FVM - finite volume method, STR - structured mesh, UNS - unstructured mesh, Y - features are available, N - features are not available

Table 3. Detailed analysis of required capabilities [19-27]

Code	Method	Thermal boundary conditions	Import			Batch mode	Temp. dependence	Phase change model
			Mesh	BC	IC			
CFX	FEM	Temperature, flux, convection, radiation, heat sourced	Y	Y	Y	Y	Y	Y
Fluent	FVM	Temperature, flux, convection, radiation, heat sourced	Y	Y	Y	Y	Y	E
Phoenics	FEM	Temperature, flux, convection, radiation	Y	Y	Y	Y	Y	E, L
FLOW-3D	FVM	Temperature, flux, convection, radiation, heat sourced	N	Y	N	Y	Y	E
Star-CD	FVM	Temperature, flux, convection, radiation	Y	Y	Y	Y	Y	E
COSMOSFlo Work	FVM	Temperature, radiation, heat sourced	N	N	N	Y	Y	N
COSMOSM	FEM	Temperature, flux, convection, radiation, heat sourced	Y	Y	Y	Y	Y	N
COSMOSWorks	FEM	Temperature, flux, convection, radiation, heat sourced	N	N	N	Y	Y	N
COSMOSDesign STAR	FEM	Temperature, flux, convection, radiation, heat sourced	N	N	N	Y	Y	N

FEM - finite element method, FVM - finite volume method, Y - features available, N - features unavailable, E - enthalpy-based phase change model, L - latent heat based phase change model

Table 4. Comparison based on possible capabilities [19-27]

CFD Comparison		**Ansys (CFX, Fluent)**	**SolidWorks (Simulation)**
Fluid Flow Capabilities	***Internal/External***	Y	Y
	Laminar/Turbulent	Y	Y
	Compressible/Incompressible	Y	Y
	Subsonic, Transonic, Supersonic	Y	Y
	Steady State/Transient	Y	Y
	Solidfication (freezing)	Y	Y
	Evaporation	Y	N
	Condensation	Y	Y
Heat Transfer Capabilities	***Conduction***	Y	Y
	Forced/Natural Convection	Y	Y
	Conjugate	Y	Y
	Radiation	Y	Y
	Solar	Y	Y
Electronics Design Capabilities	***PCB Characterizer***	N	N
	Compact Thermal Models	N	Y
	Thermostat-controlled Fans	N	Y
	Thermoelectric Coolers	N	Y
Motion Capabilities	***Motion-Driven Flow***	Y	N
	Flow-Driven Motion	Y	N
	Free Motion with Collision Detection	N	N
	Rotating / Turbomachinery	Y	Y
	Translating	Y	N
	Orbital	Y	N

Y - features available, N - features unavailable

5. Conclusion

The names and the data of the observed CFD packages, that are taken into consideration in this comparison analysis have been presented in Table 2. In some cases one company has a few codes for modelling flow and thermal problems. It is especially true with ANSYS which offers a broad range of products suited for various applications. Some of them are fully-fledged CFD packages (CFX, Fluent) while other ones are more general-purpose but offer transient heat transfer as one of the options.

The numerical method employed by the developers of packages is either FEM method or FVM. It is believed that FVM method is superior to FEM in all those applications where flows play an important role. For example, Fluent software, a FVM package, is generally recommended for flow problems by Fluent representatives. Nevertheless, there are advanced FEM-based packages that have been successfully employed in flow problems (ADINA, Abaqus, ALGOR, or Marc).

The phase change models employed in the reviewed software are based either on enthalpy or latent heat. The latent heat approach, although popular, requires caution with implementation, and especially time integration algorithms. This is due to the spike-shaped temperature/specific heat function. The enthalpy-based phase change models evaluate the effective specific heat directly from the enthalpy, and thus avoid using specific heat function with a latent heat jump.

The overall majority of the packages allow reading external files with data on mesh, initial and boundary conditions, which makes them good candidates for being used in the

coupled system. Those that do not offer one of those features could be still applied, but it would require an extensive help from a humane operator making the simulation more difficult.

In this paper we have reviewed features of several CFD packages that are mostly used nowadays both in industrial and academic research. The goal of this comparison was not to reveal a winner, but to gather the data on various codes, and show possible choices. The main goal of this paper is to be used as a guide to help end-users in their vendor selection process.

Based on all above mentioned in this paper it can be concluded that the choice of CFD software packages depends on the needs of the user and the observed problem. One of the key factor of successful usage of CFD software solutions is also the financial resources, which is also considered.

References

1. Milne-Thomson L.M., *Theoretical Aerodynamics*, Mineola, Dover Publications , N.Y., 1973.
2. Johnson N.L., *The Legacy and Future of CFD at Los Alamos*, Canadian CFD Conference, Ottawa, Canada, 1996.
3. Harlow F.H., *Fluid Dynamics in Group T-3 Los Alamos National Laboratory:(LA-UR-03-3852)*, Journal of Computational Physics, vol. 195, no. 2, pp. 414–433, 2004.
4. Harlow F.H., *A Machine Calculation Method for Hydrodynamic Problems*, Los Alamos Scientific Laboratory report LAMS, 1956.
5. Gentry R.A., R.E. Martin, J.B. Daly, *An Eulerian Differencing Method for Unsteady Compressible Flow Problems*, Journal of Computational Physics, vol. 1, no. 1, pp. 87–118, 1966.
6. Fromm J.E., F.H. Harlow, *Numerical Solution of the Problem of Vortex Street Development*, Physics of Fluids, vol. 6, pp. 975-982, 1963.
7. Harlow F.H., J.E. Welch, *Numerical Calculation of Time-Dependent Viscous Incompressible Flow of Fluid With a Free Surface*, Physics of Fluids, vol. 8, pp. 2182–2189, 1965.
8. Hess J.L., A.M.O. Smith, *Calculation of Potential Flow About Arbitrary Bodies*, Progress in Aerospace Sciences, vol. 8, pp. 1–138, 1967.
9. Rubbert P., G. Saaris, *Review and Evaluation of a Three-Dimensional Lifting Potential Flow Analysis Method for Arbitrary Configurations*, 10th Aerospace Sciences Meeting, San Diego California, pp. 72-188, 1972.
10. Al-Masry W.A., *Analysis of Bubble Column Hydrodynamics Using Computational Fluid Dynamics*, Riyadh, Kingdom of Saudi Arabia: Research Center College of Engineering King Saud University, 2006.
11. Markov D., *CFD Packages – General Features*, *Computational Engineering*, 3rd International course for young researchers, Pamporovo, Bulgaria, pp. 99-107, 2007.
12. Jones P.J., G.E. Whittle, *Computational Fluid Dynamics for Building Air Flow Prediction—Current Status and Capabilities*, Building and Environment, vol. 27, no. 3, pp. 321-338, 1992.
13. Fawcett N.S.J., *Getting Started With CFD*, Conference Computational fluid dynamics - tool or toy?, pp.1-4, 1991.
14. Shaw C.T., *Using Computational Fluid Dynamics*, Prentice Hall, N.Y., 1992.
15. Wesseling P., *Principles of Computational Fluid Dynamics*, Springer Series in Computational Mathematics, 2001.
16. Andersson B., R. Andersson, L. Hakansson, M. Mortensen, R. Sudiyo, B. Van Wachem, *Computational Fluid Dynamics for Engineers, 2nd Edition*, Cambridge University Press, 2012.
17. Chung T.J., *Computational Fluid Dynamics*, Cambridge University Press, 2002.
18. Lomax H., T. Pulliam, D. Zingg, *Fundamentals Of Computational Fluid Dynamics*, Springer, 2003.
19. Kuriyama T., *A Comparison of CFD Commercial Codes – Benchmark CFD Study of Aerodynamics and Engine Cold Flow*, Workshop on CFD in Automobile Engineering, JSAE, 1997.

20. Celuch-Marcysiak M., W.K. Gwarek, M. Sypniewski, *New Applications of the FDTD Method With Enthalpy-Dependent Media Parameters*, European Microwave Conference, London, U.K., pp. 333-336, 2001.
21. Kopyt P., W. Gwarek, *A Comparison of Commercial CFD Software Capable of Coupling to External Electromagnetic Software for Modeling of Microwave Heating Process*, The Industrial Microwave Modeling Group 6th Seminar Computer Modeling & Microwave Power Engineering, Austin, 2004.
22. Lee H.H., *Finite Element Simulations with ANSYS Workbench 14*, Mission, KS: SDC Publications, pp. 602, 2012.
23. ANSYS CFX-Pre User's Guide, Release 12.1., Canonsburg, PA: ANSYS, pp. 265, 2012.
24. Ludwig J.C., D.B. Spalding, S. Mortimore, *Starting with Phoenics-VR*, CHAM, pp. 41, 2010.
25. ANSYS Fluent 12.0, User Guide, Canonsburg, PA: ANSYS, pp. 2070, 2009.
26. SolidWorks flow Simulation, Tutorial, Massachusetts, USA: SolidWorks, pp. 240, 2009.
27. Reyes A., *Beginner's Guide to SolidWorks 2013 -Level I*, Mission, KS:SDC Publications, pp. 486, 2013.

Intelligent Systems for Control and Monitoring of Heating Systems in Individual Buildings

Milena Todorovic[1], Dragoljub Zivkovic[1], Pedja Milosavljevic[1], Dragan Pavlovic[1], Marko Mančić[1]

[1] University of Nis, Faculty of Mechanical Engineering, 18000 Nis, Serbia
milenatod1@yahoo.com

Abstract. Industry continually invests in innovative solutions for building management, including renewable energy products, heating systems, different green technologies used for lightening etc. With development of energy systems the processes of control and regulation in these systems are experiencing great progress and wide application. In field of mechanical engineering and electronics, with the problems in thermal power systems and automation have been developed new systems of regulation and energy management and accordingly development of district heating systems. Although the analysis in this paper applies to home heating systems installation, it is evident that the rational use of energy represents the strategic objectives of the energy policy. These strategic objectives, which are described with rational consumption, using alternative energy sources and environmental protection need to reconcile the growing need for energy. In this sense, the affirmation of such regulation means is inevitable. However, with technology development, particularly in the computer field, new intelligently guided systems of regulation and supervision appear, that show very good characteristics in operation. General analysis of such fuzzy logic controller is given in this paper.

Keywords. Intelligent control systems, heating systems, home installations, fuzzy controllers, green technologies.

1. Introduction

The basic issue related to energy management nowadays is relation between quality of life and rational energy consumption. The issue, in a first place related to plants, heating systems and on air-conditioning as well. The global orientation of the rational energy consumption involves many activities such as standardization, control alternative energy supplies etc. Common control systems without power supplies replaces with electromechanical, pneumatics and nowadays with electronic (analog and digital) control devices to improve working and living quality. Intelligent control express new approach to this problem, and achieve significant results [1, 2].

With energy system development the regulation process in such systems experienced great progress and wide application. If the energy systems are evaluated, that are designed in less time, with less primary energy, we will see how much energy consumers were larger but not functional. Actually, the prise of energy has caused that everyone should deal with this strategically important issue in every domain, to analyse conditions and exploitation of energy. Approaching problems in the field of mechanical engineering and electronics, as well as in the field of thermal power engineering and automation came to the development of

new system of control and regulation of thermal engineering processes, and accordingly the heating system in general.

It is imperative that systems designed today should consume less energy. However, they should meet the needs of the same both in quantitative and qualitative terms. The special attention should be paid to:

- Systems solution and mode of action;
- Thermodynamic optimum;
- Regulation;
- Proper placement of controller;
- Determining the difference between the derivative systems and projects;
- Terms of comfort and environmental factor.

Having considered the above mentioned factors, a properly regulated energy system can meet demands for energy in optimal way [3, 4].

2. Intelligent Systems

The development of system regulation, following the development of science and technology, has passed through several stages conceptually dividing systems with and without auxiliary energy (electromechanical, pneumatic and electronic systems). The electronic systems, thanks to the massive industries of electronic components, quality and efficiency, outweigh all others.

The most modern systems of control and regulation involve the use of fuzzy-logic and neural controller, which is increasingly replacing the classic analogue and digital systems.

The analysis of such fuzzy-logic controller applied to home heating installations is the subject of this paper.

Fuzzy-logic represents a method of presenting information in a way that mimics natural human communication, and thereby manipulate the information in a way characteristic to human reasoning. Fuzzy logic is applied to control systems, systems for diagnostic, monitoring, signal processing and so on. An example of generalized fuzzy logic controller is shown in Figure 1 [5, 6, 7].

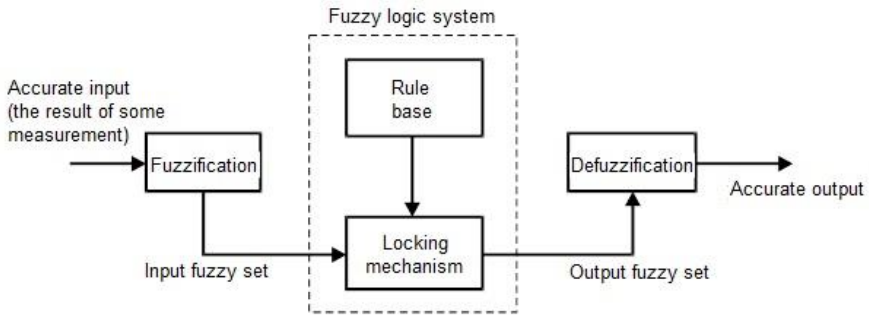


Figure 1. Fuzzy logic controller

Current trend does not involve replacing the conventional control system, but the use of fuzzy-logic as a supervisor and supplement conventional methods of regulation. Unique fuzzy-logic system is capable of to handle simultaneously numerical and linguistic skills. For many problems, we can distinguish two types of knowledge [7, 8]:

- Objective knowledge used in engineering practice;

- Subjective knowledge that represents linguistic information and it is difficult to express with classical mathematical tools.

This two types of knowledge can be coordinated by using fuzzy logic. Accordingly we have two approaches of solving specific problems where the objective information represented by a mathematical model is supplemented by subjective information that is converted into rules and then quantified by fuzzy logic.

To maximize the heat output, efficiency and comfort of home heating installations in [9] was analyzed the new generation of fuzzy logic controller for boilers. These controllers allow optimal adaptation to changing user's requirements in heating using one sensor less than in the previous generation of controllers [10]. As conventional regulators and fuzzy logic controllers are based on a standard 8-bit microprocessor, and construction, installation and optimization of fuzzy controller is supported by a software development system fuzzyTECH [11].

Most of European houses are connected to a central heating system that uses boilers for liquid fuel or gas. These boilers heat water heater and from them the hot water is distributed through pipes to radiators in rooms that are heated. In order to meet different users' needs for heat, water temperature must be constant as a function of external temperature. To measure the external temperature the external sensor is used, that is installed on the exterior facade of the building. The controller has direct impact on the start up the burners in the boiler and practically executes on-off function. For example, if feed water temperature drops two degrees below the set value, the fuel valve opens and the ignition system is started. When the temperature rises two degrees above the set value the valve is being closed [12].

Although the structure of this controller is simpler, the temperature value that should be determined as a set value, it is not as simple as it is seen. The set temperature should not be low, because in this case it cannot warm the room, on the other hand too high set temperatures causes unnecessary losses, which reflects the utilization and user comfort. Using VDI recommendations for temperature dependence, that is previously set, as a function of external temperature, represents common way of thinking that is the maximum amount of heat required to heat equal to the sum of heat losses [13]. But this point of view is not always sufficiently accurate because it excludes elements such as opening of windows and doors, ventilation, uneven heat demand during the day and the month and even the heating season and etc.

3. Fuzzy Controller of Boiler Temperature

Basically there could be two approaches for determining the set temperature, assuming proper installation and well-insulated buildings. One approach includes massive use of temperature sensors, even in every room, which affects to the system cost. The second approach implies knowledge of the data curve of the current consumption in the object, which is measured by the number of switching on and off the burners [14, 15]. An example of such curve is shown in Figure 2 and from it can be derived the four parameters:

- Current consumption (shows current load);
- The tendency of the middle member (I) (phase of increasing and decreasing heating);
- The tendency of a short member (II) (disorders, such as the opening of doors and windows);
- Yesterday's average consumption (show yesterday thermal load);

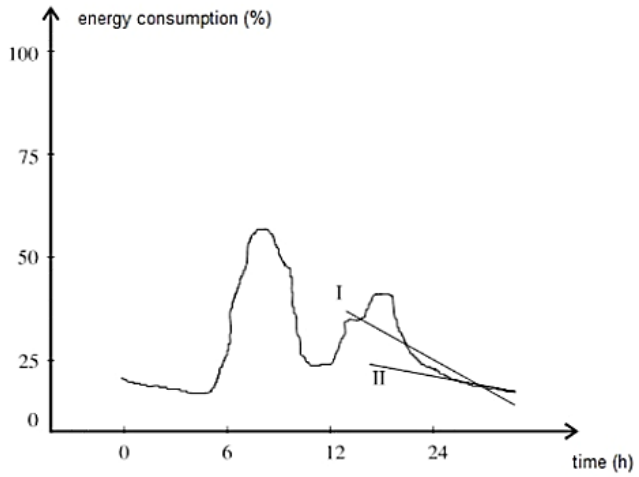


Figure 2. Current energy consumption of the object

These parameters are used to set the rules for determining the appropriate set boiler temperature. Taking into account the rules of probability, the input parameter of the system is also an outdoor temperature of the site. In this way, the installation of an external temperature sensor is unnecessary. The structure of such a controller is given in Figure 3. Picture in general represents a fuzzy controller, while a block diagram of a conventional boiler regulator is bordered with a dashed line [16, 17].

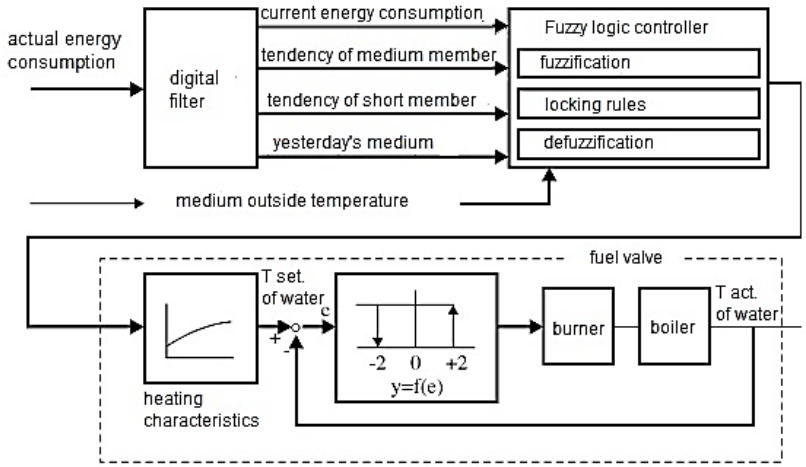


Figure 3. Schematic representation of the boiler controller

Fuzzy controller uses the sum of five pulses, which are: four pulses from the curve of energy consumption and one impulse from database of medium external temperature. The output of such system is actually estimated demand for heating energy in the house. Basically this is the main advantage and the quality of this kind of controller. The basic rule consists of "if, then" rules which could be illustrated on the example:

If - current energy consumption is low
and - the tendency of medium member is increasing
and - the tendency of short member is to falling
and - yesterday's consumption is average
and - medium outdoor temperature is very low
then - assessment of demands for heating is medium high.

In reference [9] are listed about 405 defined rules for parameter estimation. During the system optimizations some rules become more important than others. It is therefore possible association of rules with support degrees (from 0 to 1) in order to express the importance of each rule. After completing and defining the linguistic variables, functions and rules, the system is adjusted to the target hardware, which is based on the 8051 processor. Raising the level of optimization is achieved with online module, the target hardware and the development workstation (PC-Windows). This enabled the graphical representation of information in real time.

4. Comparison of Conventional and Fuzzy Controllers

In order to test the characteristics, the conventional controller and a fuzzy controller are connected to installation of the individual building [9]. Both systems have been studied for a period of 48 hours, and the result of these tests is shown in Figure 4.

Three curves of temperature change in time for system test are represented. Two curves are related to tested system of regulation including: water temperature produced with the classic controller according to the outside temperature and the temperature obtained with the fuzzy controller. The third temperature represents the optimal water temperature calculated from the external and internal conditions of the building.

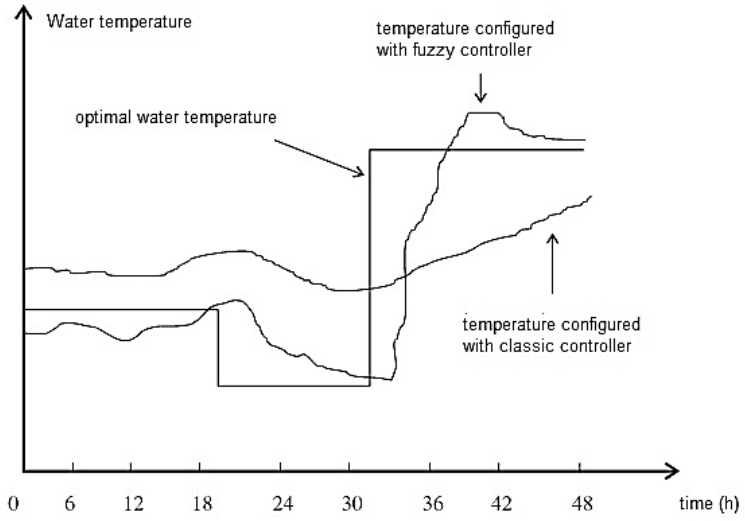


Figure 4. Comparison of temperature characteristics

The result of the comparison shows that the fuzzy controller is very well suited to the current needs for heating. By setting the temperature below the level used by the conventional controller, especially during periods of low load, this system effectively saves

energy which is evident in longer periods of use. Besides, because of the reduction of energy production heat utilization is improved. It is observed that heat comfort is increasing especially in terms of sudden heat demands, the values can be easily set, and savings in manufacturing and installation costs are evident.

5. Conclusion

Undoubtedly the characteristics of new regulation modes have certain advantages over conventional systems. Although the analysis of this paper refers to the system of home heating installations, it is evident that the rational consumption of energy in this micro plan can be achieved the strategic objectives of energy policy. These strategic objectives, described with rational consumption, the application of alternative energy sources and environmental protection need to reconcile the growing need for energy. In this sense, the affirmation of such kind of regulation is inevitable.

References

1. Berenji H.R., Fuzzy Logic Controllers, The International Series in Engineering and Computer Science, vol. 165, pp. 69-96, 1992.
2. Chen Y.Y., *Stability Analysis of Fuzzy Control—A Lyapunov Approach*, Conference IEEE Systems Man Cybernetics, volume 3, pp. 1027–1031, 1987.
3. Jian Z., S. He-Xu, Z. Jiang-tao, *Application of Adaptive Fuzzy Sliding-Mode Controller for Heat Exchanger System in District Heating*, International Conference Intelligent Computation Technology and Automation (ICICTA), vol. 1, pp. 850-854, 2008.
4. Xu B., L. Fu, H. Di, *Field Investigation on Consumer Behaviour and Hydraulic Performance of a District Heating System in Tianjin, China*, Building and Environment, vol. 44, no. 2, pp. 249-259, 2009.
5. Zadeh L.A., *Fuzzy Sets*, Information and Control, vol. 8, pp. 338–353, 1965.
6. Zadeh L.A., *A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges*, Journal of Cybernetics, vol. 2, pp. 4–34, 1972.
7. Mamdani E.H., S. Assilian, *An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller*, International Journal of Man-Machine Studies, vol. 7, no. 1, pp. 1–13, 1975.
8. Lee C.C., *Fuzzy Logic in Control Systems: Fuzzy Logic Controller*, Systems, Man and Cybernetics, IEEE Transactions, vol. 20, no. 2, pp. 404-418, 1990.
9. Altrock C.V., H.O. Arend, B. Krause, C. Steftens, E. Behrens-Rommel, *Adaptive Fuzzy Control Applied to Home Heating System*, Fuzzy Sets and systems, vol. 61, pp. 29-35, 1994.
10. Mendel J.M., *Fuzzy Logic Systems for Engineering*, Conference IEEE, vol. 83, no. 3, pp. 345-377, 1995.
11. Kickert W.J.M., E.H. Mamdani, *Analysis of a Fuzzy Logic Controller*, Fuzzy Sets and Systems, vol. 1, no. 1, pp. 29-44, 1978.
12. Jang J.S.R., C.T. Sun, *Neuro-Fuzzy Modeling and Control*, Conference IEEE, vol. 83, no. 3, pp. 378-406, 1995.
13. Goli D.K, *Automation of Air Conditioning Unit*, SMEITS, pp. 427, 1989.
14. Vilar J.R., C. Enrique de la, J. Sedano, *A Fuzzy Logic Based Efficient Energy Saving Approach for Domestic Heating Systems*, Integrated Computer-Aided Engineering, vol. 16, no. 2, pp. 151-163, 2009.
15. Prokhorenkov A.M., A.S. Sovlukov, *Fuzzy Models in Control Systems of Boiler Aggregate Technological Processes*, Computer Standards & Interfaces, vol. 24, no. 2, pp. 151-159, 2002.
16. Lianzhong L., M. Zaheeruddin, *Hybrid Fuzzy Logic Control Strategies for Hot Water District Heating Systems*, Building Services Engineering Research and Technology, vol. 28, no. 1, pp. 35-53, 2007.
17. Clarke J.A., J. Cockroft, S. Conner, J.W. Hand, N.J. Kelly, R. Moore, T. O'Brien, P. Strachan, *Simulation-Assisted Control in Building Energy Management Systems*, Energy and Buildings, vol. 34, no. 9, pp. 933-940, 2002.

NuOffice Munich – A Role Model for Sustainable Buildings

Milena Todorovic[1], Dragan Pavlovic[1], Dragoljub Zivkovic[1], Pedja Milosavljevic[1]

[1] University of Nis, Faculty of Mechanical Engineering, 18000 Nis, Serbia
milenatod1@yahoo.com

Absract. In order to raise the limit of the building's energy efficiency, the science and industry came together to maximize the building's passive energy use. It can be achieved by driving down energy consumption and introducing renewable energy sources as main suppliers to the building. One of the successful projects that has been implemented so far is in Munich. Germany now have the greenest office building in the world. The NuOffice, one of the three DIRECTION demonstration sites, has been awarded the certificate for Leadership in Energy and Environment Design (LEED) and was assigned LEED platinum status. This business facility can easily serve as a model for future new construction across the globe, making not only savings in energy consumptions but also the impact on environmental green thinking. This paper includes all available data regarding this project, goals and achieved results. It is very important to have complete insight into the latest news in the green projects and energy efficient objects. This project as a role model can help new idea to come and to have such projects in our country.

Keywords. Passive buildings, green technologies, solar energy, energy saving, LEED certification.

1. Introduction

Most of the people spend the majority of their time at work. Whether it's a traditional office building with separate departments, individual offices or conference rooms, or co-working space, the workspace is definitely one of the important part of lives of working people. It should be treated also as a healthy place. No matter how green nowadays offices may be, it's likely to pale in comparison to the NuOffice, however.

Recently completed in Munich, NuOffice represents commercial property. It is considered to be the greenest office building in the world. Commissioned by Haupt Immobilien, and created with the help of research group DIRECTION and the Fraunhofer Institute for Building Physics, NuOffice recently received the highest Leadership in Energy and Environment Design (LEED) rating ever issued for a building of its type. The US Green Building Council (USGBC), which manages the LEED green building certification program, awarded NuOffice an 94 points. This rating is 80 points that are required to achieve LEED Platinum, currently the highest standard for an environmentally-friendly structure recognized by the USGBC [1].

This paper presents some part of this project, as well as the investigation that preceded it's realization. In order to have more green objects in our environment, it is very important to understand what it can be done with resources that are at our disposal and to be not only environmentally friendly but also to have more economic approach to our living.

2. Photovoltaic and Zero Energy Building

According to the European Directive 2010/31/EU, starting from the end of 2020, all new buildings will have to be Nearly Zero Energy Buildings (Nearly ZEBs). According to this directive, 'Nearly ZEB' means a building that has a very low energy yearly energy consumption, which can be achieved by both the highest energy efficiency and by energy from renewable sources, which shall be 'on-site' or 'nearby' [2].

In the near future, buildings will be designed to have a need for very little energy (passive design strategies for energy efficiency) and to integrate active surfaces (i.e. PV modules) for energy generation. This approach requires new point of view that will use the energy that is needed as an input that will influence on design. The usable energy should be seen as a variable able to relate itself to the form of buildings (or clusters of buildings or even cities and landscapes), instead of being seen as a kind of abstract variable that design cannot deal with. In the future, from designers will be expected to consider not only the space that is used directly but also the space required to provide generation of electrical and thermal energy from renewable sources: for example the surface necessary for placing the energy generation devices. This area can be defined as the 'building's energy footprint' [3]. The problem will be the fact that the renewable energy generation systems are visible and demand bigger space, comparing to the to conventional energy sources. It can be also presented as a challenge, because for the first time in the tradition of architecture, energy source can represent a 'form' (i.e. shape, colors and features of a PV generator), and architects will be responsible for designing this form.

Because of the high energy consumption of the European countries, PV can contribute significantly to the reduction of the primary, conventional energy supply, as well as to the reduction of the CO2 emissions [4]. PV seems to be technically the easiest way to obtain the zero energy balance, as well as, drop in prices makes it competitive even with active solar thermal collectors and building materials in general. But, there is limitation of energy quantity production per square meter PV collector (it depends mostly on the PV energy efficiency, on the tilt and azimuth angles of the PV generator and on the latitude, as well as on the BOS efficiency). As an example, in the case of optimal positioning of the PV system (tilt and azimuth angles), a typical generation in northern climates such as North Europe would be around 80–100 kWh/m2 collectors per year as usable energy. In southern climates such as South Europe, a typical generation would be around 130–180 kWh/m2 collectors per year as usable energy [5].

Finally, it is important to say that, the use of PV in buildings is under investigation since more than 20 years now and it has been recognized as a key factor for the exploitation of PV and the reduction of the CO2 emissions of buildings, making not only the buildings greener but also the living environment [6, 7, 8, 9].

3. About the Project NuOffice

The NuOffice in Munich is a project of the Hubert Haupt Immobilien Holding. In collaboration with the Fraunhofer Institute and Munich Technical University, new prototypes of efficient office properties are being developed in three construction phases. It is planned to be completed by 2015 in Parkstadt Schwabing, with a total area of approximately 33,000 m². The first building phase was completed in December 2012. Project developer Hubert Haupt had the goal right from the start that only the Platinum LEED certificate would meet requirements of this project. One of the statements that Haupt gave was that the only goal of this building was to comply the government goal for the years 2050 to 2100. This means that

the energy standard that will only apply 40-90 years from now has already been exceeded. As an example of this statement is can be consider the limit of CO2 emissions. Groundwater is used for heating and cooling purposes, while smart lighting and ventilation concepts ensure a low consumption of energy, which itself is generated by the photovoltaic system installed on the roof of the building [10].

On July 1, 2013 the NuOffice in Munich was awarded the LEED certification in "platinum" – the highest acknowledgement given by the certification system developed by the U.S. Green Building Council for ecological building. With the best ever score of 94 points awarded in the category "Core and Shell" the NuOffice has positioned itself as the most sustainable office building in the world. Core and Shell concerns the entire base building, including mechanical, electrical, plumbing, and fire protection systems, and it is typically used for commercial buildings of this nature.

The team members responsible for the demonstration activities in Munich are DIRECTION partners Domagk Gewerbepark, FACIT and Fraunhofer IBP.

Research group that participated in the project creation is named DIRECTION, and this group is co-financed by the EU within Framework Programme 7 (FP7). DIRECTION is coordinated by Fundación CARTIF (Boecillo – Valladolid) and developed in cooperation with other 10 partners: DRAGADOS and 1A Ingenieros ; EURAC, the Autonomous Province of Bolzano and Claudio Lucchin & Architetti Associati; Domagk Gewerbepark, Fraunhofer Institut für Bauphysik and FACIT; Enginsoft SpA ; youris.com G.E.I.E. [11].

This project was commissioned by Hubert Haupt Immobilien Holding, that is founded in 1994 and managed by its owner Hubert Haupt. Hubert Haupt Immobilien Holding specializes in commercial and residential building construction. Facit GmbH & Co. KG, which is part of the Holding, is responsible of project planning while marketing and leasing activities are carried out by the subsidiary Hubert Haupt Immobilien management GmbH. The company has realized more than 40 large projects so far [12].

4. Standards in Green Construction - LEED Certification

In 1990, the Government of United Kingdom pioneered the green standards at request of the British real estate industry. The BREEAM – the Building Research Establishment's Environmental Assessment Method was launched. BREEAM evaluates the environmental performance of a broad spectrum of new and existing UK buildings [13].

The creation of reliable building-rating and performance measurement systems for new construction and renovation has helped the change of corporative perceptions about green thinking [14]. In 2000 the US Green Building Council (USGBC) in Washington, introduced its rigorous Leadership in Energy and Environmental Design (LEED) building rating system [15]. It was recognized just as one of several independent systems for rating "green buildings," today it has emerged as the leading green building rating system. Since its inception, different approaches of analysis have been done to meet the statement that the LEED buildings are in substantially the energy-efficient buildings, and also have been supported by data [16, 17].

LEED evaluates building and awards points in six areas, such as innovation and design process. The program awards points in the following categories: sustainable site (14 possible points), water efficiency (5 possible points), energy atmosphere (17 possible points), materials and resources (13 possible points), indoor environmental quality (15 possible points), and innovation and design process (5 possible points). Companies can earn points for everything from brownfield redevelopment to public transportation access. LEED has four award levels: Certified (26-32 points), Silver (33-38 points), Gold (39-51 points) and

Platinum (52-69 points) [15]. A LEED-Gold building has 50% less negative impact on the environment than standard building. A LEED-Platinum building has at least 70% less negative impact [14].

More and more countries are creating their own green standards. The Green Building Council of Australia, founded in 2002, synthesized BREEAM, LEED and other environmental criteria into the Green Star rating system, which is specific to the Australian environment, building practices and real estate markets [18]. India's Green Building Council is developing a rating system, which was launched in 2007 [19].

One of the advantages is that the certification assures prospective buyers and tenants that the building is truly sustainable [13]. Hundreds of U.S. and international studies have proven also the financial advantages of the green buildings. Well-designed green building, as it is known, have lower utility costs. In its first year of operation, Genzyme Center – Genzyme Corporation's the LEED Platinum headquarters in Cambridge, Massachusetts – used 42% less energy and 34% less water than standard buildings of comparable size [21]. Green buildings can also boost employee productivity by approximately 15%, in part because they use alternative building materials that don't have toxic emissions, like formaldehyde, that are commonly found in standard building materials and work-places. What is also important that the green design criteria – including abundant day-lighting, individual climate controls and outdoor views – raise morale and employee satisfaction, which also improves productivity [22].

What is also important to be mentioned is that the green building materials, mechanical systems and furnishing have become more widely available, and their prices have dropped considerably. According to Turner Construction chairman Thomas C. Leppert, four industry studies of more than 150 sustainable buildings across the United States show that, on average, it costs only 0.8% more to achieve basic LEED certification than to construct a standard building [23].

LEED certification employs a variety of assessment systems for different buildings, and it is against this background that a separate system was developed for new buildings used by third parties, known as LEED Core & Shell; this is the system by which the NuOffice building in the Domagkstrasse in Munich was tested and categorized.

An innovative energy concept, developed by experts of the Fraunhoferinstitut, served as the basis for achieving LEED Platinum, scoring 37 out of 37 possible points in the category "Energy and Atmosphere". An absorption heat pump making use of the well water's heat potential and bringing it up to heating temperature stands as a true innovation on the German market. The system runs on district heat, which has a favourable primary energy factor, thus eliminating the need for an electric heat pump. A photovoltaic system, covering nearly the entire roof, feeds electric power into the building and into the grid. The optimizing of all energy-related systems during the planning and construction phase proved to be especially valuable during the commissioning phase.

In addition to its energy performance, the NuOffice building also scored with regard to its location: Offering excellent access to public transportation and local amenities, a favourable bi-cycle infrastructure and an on-site electric vehicle charging station, the building achieved 25 out of 28 points. Ecological criteria such as abundant green spaces and 100% storm-water infiltration complemented the score sheet. Targeted lighting minimizes "light pollution"; water consumption is reduced by efficient fittings in all rental spaces while outside spaces are completely independent from artificial irrigation.

Recycling rate of 84% construction waste, such as regional building materials, FSC certified wood and sustainable user qualities are among the key factors contributing to the

record score. Exemplary measures include CO2 monitoring in the rental spaces, under-pressure in all chemically contaminated rooms such as cleaning agent and waste rooms or the parking garage as well as relatively simple concepts such as doormats in front of all entrances reducing dirt within the building.

The LEED category "Innovation in Design" recognizes over-achievement of requirements as well as outstanding and innovative solutions. The NuOffice building's projected energy savings of 53% compared with the reference case fall within this category as well as the "Green Building Monitor" installed in the entrance area indicating current energy and water consumption as well as photovoltaic power generation.

5. Building Properties

Turning away from the glass-walled style of office buildings that have been popular in recent years, the windows at NuOffice are more modest in size and triple-glazed. This means cooler interiors on hot days, reducing the need for air conditioning and less heating on cooler days. Sensors that are installed, measure the temperature and air humidity both inside and outside the building. This data is used to regulate the indoor climate [1, 10].

Essential concept components are targeted on use of environmentally compatible building material, high quality of ambient air, maximum flexibility of occupancy and high energy efficiency for heating, cooling, ventilating and lightening.

This includes excellent heat insulation, innovative systems engineering, use of renewable energy, passive heating and cooling concepts as well as optimized daylight concepts and artificial light. The core of the innovative energy concept is the application of an absorption heat pump, operated by district heat having a certificated and extremely advantageous primary energy factor of 0.122. As low temperature heat source, heat is extracted from ground water using an adaptor to a well. With this concept, the use of district heating in Munich will be further optimized.

The building will be equipped with a surface heating and cooling system designed for thermal activated building constructions. During the heating period, energy supplied by the absorption heat pump will be used while in cooling periods, a free cooling system using ground water will be realized. However, in peak seasons, the cooling capacity will be also actively supported by the operation of an absorption heat pump.

Basic requirement for the specified energy supply concept which has never been realized in this form, is the development respectively adjustment of a heat pump to work both with district heat with low supply and low return temperatures. This applies both for heating and cooling, for which high efficiencies and suitable temperature levels need to be maintained even with unusual operating conditions. Detailed and integrative planning of the entire energy concept is necessary for the design of the absorption heat pump as well as the system and its operation.

Energy efficiency solutions applied:

- Improved insulation of exterior walls will be realized with 30 cm polystyrene instead of the conventional insulation systems with 14 cm. Improved thickness in roof constructions will be also installed.
- Triple glazing windows with high efficient framing.
- External and internal shading-systems for west and east-facing offices, electrochrome glass in south-oriented facade for improved overheating protection.
- Using district heat in combination with an absorption heat-pump having an overall efficiency of 1.53.
- Use of groundwater for passive cooling.

• Demand controlled ventilation system with semi-decentralized ventilation units that permit high heat recoveries.

• High efficient lighting using LED technology.

• Demand controlled artificial lighting management considering occupancy and daylight illumination.

With all the achieved savings in operating costs Domagk - Gewerbepark is able to offer its tenants a utility costs flat rate, which is unique on the private rental market in Munich.

6. Future Projects

A new project has been launched end of January 2012 trying to find answers to the question: How to enhance the overall energy efficiency of a building in order to achieve a consumption level of primary energy lower than 60 kWh per m2 per year? A new project has been launched in the end of January and it will be co-financed by the EU within Framework Programme 7 (FP7). During the next four years, building and construction engineers, architects, energy researchers, IT specialists and public authorities will work together in order to show on three different sites in Europe how the ambitious goal can be reached. The sites are located in Valladolid (ES), Bolzano (IT) and Munich (DE). The project is called DIRECTION which stands for "Demonstration of very low energy new buildings".

Based on the analysis of suitable energy efficiency technologies and their technical and economic viability, the demonstration activity will be deployed at the sites in three new buildings. In each of the three buildings a set of very innovative measures such as constructive elements for energy optimization, high efficient energy equipment and advanced energy management will be applied.

DIRECTION is set to provoke significant impacts for new buildings. The building sector business will be boosted for very low energy new buildings. Energy consumption reductions of more than 50% and CO2 emission reductions of more than 60% are expected. New standards and regulations for low energy buildings need to be implemented by European and national policy-makers. Local, national and European stakeholders including public authorities, users and citizens at large will be kept up-to-date about the progress and the outcomes of the demonstration [11].

The project is managed by the innovation centre Fundación CARTIF (Boecillo – Valladolid) who is also responsible for the demonstration activities at the Valladolid site. Other partners of the project include DRAGADOS and 1A Ingenieros , both located in Spain; EURAC, the Province of Bolzano and Claudio Lucchin & architetti associate, all located in Italy and responsible for the Bolzano demonstration site; Domagk Gewerbepark, Fraunhofer Institut für Bauphysik and FACIT GmbH und Co. KG, all dealing with the demonstration site in Germany. The partnership is complemented by Enginsoft SpA (IT) dealing with software development and youris.com G.E.I.E (BE) in charge of dissemination and communication.

7. Conclusion

One of the negative consequences that follow the development of the industry in the world is the increased pollution of environment. This is the reason that today; one of the crucial issues for humanity is how to preserve the environment and to reduce a greenhouse effect.

There are plenty of examples of "Green Buildings" in the world and their savings in energy consumption are mostly specified and some of them have reached LEED Certification status. The Nu-Office is an office building, located in Domagkstrasse. From an eco-friendly point of view the Nu-Office is built to "Sustainable Building" standards and a unique example in Germany. This building meets very high energy efficiency standards and maximum comfort requirements.

Commissioned by Haupt Immobilien, and created with the help of both European-funded research group DIRECTION and the Fraunhofer Institute for Building Physics, NuOffice breezed through LEED Platinum certification. It snagged the highest rating ever issued by the body for a building of its type. The LEED Platinum certificate was handed over on July 1st, 2013. This project presents the latest installed achievements in this area. This is the also very good example of merging the environmental and economic benefits.

References

1. http://www.usgbc.org/projects/nuoffice-i
2. EPBD recast, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast), Official Journal of the European Union, 2010, (http://eur-lex.europa.eu/LexUriServ/ LexUriServ.do?uri=OJ:L:2010:153:0013:0035:EN:PDF).
3. Scognamiglio A., H. Ossenbrink, M. Annunziato, *Forms of Cities for Energy Self-Sufficiency*, The 24th World Congress of Architecture, pp. 22–27, 2011.
4. Voss K., K. Kiefer, C. Reise, T. Meyer, *Building Energy Concepts with Photovoltaics—Concept And Examples from Germany*, Advances in solar energy, American Solar Energy Society, Boulder, vol.15, pp. 235–259, 2003.
5. http://re.jrc.ec.europa.eu/pvgis/apps4/pvest.php
6. Abbate C., *L'integrazione Architettonica Del Fotovoltaico, Casi Studio Internazionali Della Tecnologia Solare*, Gangemi Editore, pp. 256, 2006.
7. Hagemann I.B., *Gebäudeintegrierte Photovoltaik, Architektonische Integration Der Photovoltaik in Die Gebäudehülle*, Verlagsges Müller, pp. 433, 2002.
8. Humm O., P. Toggweiler, *Photovoltaik und Architektur—Die Integration Von Solarzellen in Gebäudehüllen*, Birkhäuser Verlag, pp. 120, 1993.
9. Schoen T.J.N., H. Schmidt, *Building-integrated PV Activities within the IEA Framework: Going from Task 16 to Task 7*, Progress in Photovoltaics: Research and Applications, vol.6, pp. 237–245, 1996.
10. Gaiddon B., H. Kaan, D. Munro, *Photovoltaics in the Urban Environment*, Lessons Learnt from Large-scale Projects, Routledge, pp. 208, 2009.
11. http://www.nuoffice.de/
12. http://www.direction-fp7.eu/
13. http://www.haupt-immobilien.de/
14. Lee W.L., J. Burnett, *Benchmarking Energy use Assessment of HK-BEAM, BREEAM and LEED*, Building and Environment, vol.43, pp. 1882-1891, 2008.
15. Lockwood C., *Building the Green Way, Tool Kit*, Harvard Business Review, pp. 9, 2006.
16. Diamond R., M. Opitz, T. Hicks, B. Von Neida, S. Herrera, *Evaluating the Energy Performance of the First Generation of LEED-Certified Commercial Buildings*, The 2006 ACEEE summer study on energy efficiency in buildings, vol.3, pp. 41-52, 2006.
17. http://www.usgbc.org/
18. Turner C., M. Frankel, *Energy Performance of LEED for New Construction Buildings – Final Report*, New Buildings Institute, pp. 46, 2008.
19. Newell G., *The Strategic Significance of Environmental Sustainability by Australian-Listed Property Trusts*, Journal of Property Investment and Finance, vol.26, pp. 522-540, 2008.
20. Potbhare V., M. Syal, M. Arif, M.A. Khalfan, M. Egbu, *Emergence of Green Building Guidelines in Developed Countries and Their Impact in India*, Journal of Engineering, Design and Technology, vol.7, pp. 99-121, 2009.
21. Scofield J., *A Re-examination of the NBI LEED Building Energy Consumption Study*, International Energy Program Evaluation Conference, 2009.
22. Ng B.H., Z.A. Akasah, *Reliability and Validity of a New Energy-Efficient Building Performance Survey Framework in Hot and Humid Climatic Regions: Malesia*, International Journal of Renewable Energy Research, vol.2, pp. 702-711, 2012.

Software for Energy System and Building Simulation: A Review

Marko Mančić[1], Dragoljub Živković[1], Pedja Milosavljević[1], Milena Todorović[1]

[1] University in Niš, Mechanical Engineering Faculty in Niš, Serbia, Aleksandra Medvedeva 14 Street, Niš, Serbia
markomancic@yahoo.com

Absract. Constant increase in energy consumption and demand, followed by the inevitable green house gas emissions conveys new challenges in the terms of energy efficiency and sustainable development. Global challenge of sustainable development implies increased renewable energy utilization, improving the efficiency of energy conversion, lower gas emissions and reduced dependency on fossil fuels. Systematic approach with integration of processes and technologies into polygeneration systems reveals overall benefits of combined technologies, and pinpoints integration conditions and possible problems. Fulfillment of these conditions is followed by a substantial use of financial resources. Due to complicated structure and complex influences between parameters which define the techno-economic operating conditions of these systems their optimization is a necessity. Until now, many software with abilities for different simulations and optimizations of energy systems have been developed. The objective of this paper is to determine and compare features of these software, thus pinpointing tools most suitable for the mentioned challenges of energy system design and integration. For this purpose, a comparative analysis of available energy system software is carried out and presented. Features, advantages/disadvantages are discussed and remarks about features, inputs/outputs, technologies included and possibilities for integration and system optimization are made. As a result, software with most tools and most flexibility are identified.

Keywords. Energy system simulation, polygeneration, optimization, process integration, software.

1. Introduction

A growing energy demand, green house gas (GHG) emission reductions and energy efficiency improvement are common issues of sustainable development. These sustainable development challenges, influenced by limited resources of fossil fuels led to promotion of using renewable energy sources (RES) in the highest possible extent, often with targets set by government bodies and/or European directives. However, challenged by the mentioned global trends, efforts are made towards higher efficiencies, higher RES utilization ratio and lower GHG emission in energy production. The transition problem is expanded even further by adding factors such as: dependence of fossil fuels, availability of RES, growing energy demand, energy market prices and other economical factors.

A logical and effective way through the transition is process integration aimed to provide better overall energy efficiency by combining advantages of available technologies to meet energy demands. For design, simulation and optimization of energy systems, many software were developed in the last couple of decades. Different approaches were used for

the design of these programs, making them suitable for different purposes, results and taking into account different level of input data detail and affecting the precision of the results. In order to save processing time, two types of solvers can generally be found: sequential where the calculation is carried out for example one component at the time, and simultaneous, where all the equations during simulation are solved at the same time. Elmegaard conducted an analysis of software for power plant process simulation and reported it was too specific for a narrow range of problems, difficult to learn, not able to handle model problems robustly, not very well documented in literature and/or very expensive [1,2] and analyzed the robustness of components in the energy system simulators [3]. A comparison of software for CHP process simulations with different problem approaches (dynamic or static) and sequential and/or simultaneous solvers used in the software's code are presented in [4] and made a comparison of the simulation results of a real CHP facility using three different simulators can be found in [5]. Building simulation software was compared by Crawley et. al. in 2005 [6]. Many software developers have made their software more flexible and included the possibility for design, simulation and/or optimization of hybrid energy systems, RES utilization and availability, techno-economic factors and more.

This paper gives the analysis of the energy system simulation software, with an energy system defined as any system capable to satisfy one or more energy demands. The scope of this paper is to give brief descriptions about the energy system software, their similarities, differences, features, capabilities, limitations, and help modelers to choose a most suitable package for different tasks.

2. Methodology

A broad range of software applications has been developed for modeling and simulation of energy systems and processes. Generally speaking, any tool capable of solving a given system of equations may be considered a simulation tool. Nevertheless, more specialized and more user friendly applications have been developed, and are available as free applications and/or commercial software.

A brief survey of the energy system simulation software is presented below. The survey contains data about the general solvers, building simulation software, hybrid energy system simulation software, CHP and thermal power plant simulation software. Information about the available software was gathered mostly from the available official software documentation. The survey focuses and tries to provide data regarding most important features of the software such as:

1. Area/purpose the software was designed for – the criteria will make a classification of programs suitable for different modeling applications – energy system modeling, design, optimization, building energy efficiency, process integration.
2. Type of simulation- static/dynamic
3. Solver type: simultaneous/sequential
4. Included tools and features: design, integration tools, optimization tools, LCA, techno-economic analysis
5. Included model libraries of available technologies – a classification based on the technologies covered in the ground file of the software (i.e. wind, solar, heat pump, cogeneration, gas turbine, oto/diesel engines…)
6. Flexibility – ability to combine technologies into a hybrid system, ability to make changes to the available component models.
7. Decision making- input/output options and results
8. Program code open to changes and interaction

9. User interface – Graphical user interface (GUI) or command line

3. Available Simulation And Optimization Software

BEopt is a program designed for design and optimization of zero net energy buildings: grid-tied, net-metered PV and active solar to produce as much energy as it uses on an annual source energy basis [7]. BEopt uses DOE-2 and TRNSYS simulation engines. The DOE-2 simulation program is used to calculate energy use as a function of building-envelope options and HVAC equipment options. Appliance and lighting option energy savings are calculated based on energy-use intensity factors and schedules input into DOE2. The TRNSYS simulation program is used to calculate water-heating loads and energy savings for solar water heating. TRNSYS is also used to calculate annual electrical energy production from a grid-tied PV system. For each building design, the browser will display detailed results regarding energy consumption, costs, and options. If multiple cases exist in a project file, a Combined Graphs output screen will be available.

BLAST is a building loads analysis and system thermodynamics (BLAST) tool, developed by the US Army Construction Engineering Research Laboratory, USA, CERL [6]. It has a set of programs for predicting energy consumption and both energy system performance and energy costs in buildings. Blast contains three major features: Space Load Prediction, Air System Simulation and Central Plant. The Space load Simulator uses the computed space loads, weather data and user described air system to calculate zone loads and translate them to hot/chilled water, steam, and electric building demands for a central plant system. Central plant simulation can include chillers, on-site power generating equipment and solar energy modules in the simulation to determine final buildings consumptions. The calculations are done using basic heat and mass balance equations in the programs simulator. New or retrofit building energy performance with an annual consumption estimation can be calculated. BLAST's development has been stopped in 1998, when it was merged with DOE2. It has a sequential solver [6,8]

DOE-2 predicts hourly energy use and energy costs of buildings taking into account hourly weather data, building geometric description, HVAC and utility structure [6]. It is intended for decision making about cost effective building parameters which improve energy efficiency while maintaining thermal comfort. It consists of the following tools/features: LOADS – a simulation subprogram for calculating sensible and latent components of hurly heating/cooling loads for each constant temperature taking into account weather data and building use parameters; SYSTEMS subprogram handles secondary systems, calculating performance of air equipment (fans, coils, ducts) and corrects the constant temperature calculated by the LOAD subprogram, and PLANT calculates primary systems (simulates the behavior of boilers, chillers, cooling towers, storage tanks etc, taking into account part-load of the primary equipment for fuel and electrical demand of the building. ECONOMICS subprogram calculates costs of energy, and cost benefits of different designs or savings for retrofits. DOE2 has a sequential solver, and its engine has been widely used by third party developers who created more than 20 interfaces which made the program easier to use [6,8]

gPROMS [9] is a platform for high-fidelity predictive modelling for the process industries developed by Process System Enterprise (PSE). It consists of gProms model builder supporting a wide range of functions such as: model development from first principals or existing model libraries, validation of models against experimental data using parameter estimation and experimental design, use of models for model based activities i.e. steady state and dynamic simulation and optimization. Models can be exported using CAPE OPEN tool [10] for further use with other PSE tools allowing gProms model to be

incorporated as steady state unit operation within a flowsheet tool such as ASPEn or Hysis; or incorporated in mathworks or simulink where continuous time and discrete time are supported; or exporting models into mathworks mathlab.

RETScreen Clean Energy Project Analysis Software [11] is a free-of-charge software for evaluating the energy production and savings, costs, emission reductions, financial viability and risk for various types of Renewable-energy and Energy-efficient Technologies available in many languages. Software development was supported by the Ministry of Natural Resources of Canada. The software is designed to make comparison between for example conventional "base case" technologies and clean energy "proposed case" technologies. It features cost analysis, GHG analysis, sensitivity and risk analysis based on the Monte Carlo method and a financial summary for the given energy model. Energy models included are: wind energy, small hydro plants, photovoltaic, biomass heating, solar air heating, solar water heating, passive solar heating, ground source heat pump and CHP.

INSEL is a general-purpose graphical programming language, which can generally solve any problem of computer simulation. Its has a descriptive code, and is open for adding additional user specified models. The basic version of INSEL has been developed during the period 1986 to 1991 at the German University of Oldenburg in a project of the former Federal Ministry of Research and Technology BMFT [12,13, 37]. INSEL features modules for simulation solar irradiance simulation, photovoltaics and solar thermal applications. Highly efficient stochaistical algorithms can generate time series in hourly resolution from the monthly data for one or more years, with the correct year-to-year variability of the synthetic data. Software has weather data and can simulate market available PV modules and inverters, solar heating and cooling, energy storages with fully mixed and stratified tank modules, several models of heat exchangers and solar thermal cooling with absorption technology or open air-based sorption systems, but also solar power plants (parabolic concentrating with direct evaporation and thermo oil circuts, solar tower systems, Dish- stirling Machines and up wind power plants. The software has flexibility of creating system models and configuration which is both useful and interesting feature. However, this software is intended just for simulation, not optimization, but it is free of charge [37].

TRNSYS is software for transient simulation and optimization of energy systems, developed at the University of Wisconsin, USA. [15-19, 37]. It has a complete extendable simulation environment for transient system condition simulations, as well as multi zone buildings. It is a tool aimed for validation of new energy concepts, from simple systems through buildings and their equipment as well as determination of strategies for control and alternative energy systems, weather, thermal and hydro storage... It has an open modular structure which enables adding custom/user made component models by using standard program languages. It can also be connected to other applications for post processing or interactive connection during simulation. TRNSYS consists of the following component modules: Solar systems (thermal and PV), Zero energy buildings and HVAC, Renewable energy sources, Cogeneration and fuel cells, and generally it's engine is capable of simulating anything requiring a dynamic simulation. It has tri sub-applications: Simulation Studio, TRNBuild – interface for building data input, EDITOR used for creating self executable applications for further distribution. A frame of a new component or modification of existing components can be created/modified using FORTRAN wizard. Apart from a large number of included components and technologies, and additional component libraries available online, its open structure makes it a very useful, flexible and powerful dynamic

simulation tool. It is referred to as the most complete solar energy system modeling and simulation software in [16], although it is capable of precise dynamic simulation of a far larger extent of energy systems.

ENERGY PRO Energy pro is a comercial modeling software package for techno-economic modeling, analysis and optimization in complex energy projects with a combined supply/demand of electricity, thermal energy – steam, hot water or cooling using multiple production units to meet the demands [20]. It covers a wide range of technologies for the production of heat and electricity. It consists of models of CHP, gas engines, boilers, thermal storage, industrial cogeneration plants for the production of heat, electricity and steam and hot water, absorption and electric chillers for trigeneration, biogas and other types of fuels, but also wind plants and solar thermal and photovoltaics as renewable energy sources. The program gives time-step simulations with different time steps and time function performance curves for different technologies. Factors such as priority of usage of a given technology (production unit), market prices, performance time curves and so on can be accounted for and simulated in the program by using weighing factors. Great attention has been made towards economic factors, prices tariffs, taxes and so on and the optimization can be don according to these factors. It can also make calculations of CO2 and NOx emissions, heat, cold and electricity storage, electricity markets with different tariff systems. All components in the simulator, except for economics and finance, and solar and wind farms, are modeled as black box, with input and output flows from a virtual border made around the component. The solar and wind simulations are done with more detail, based on weather data and other relevant parameters. The solver is capable of annual simultaneous simulations with constant time steps and values kept constant during these time steps. Simulation can be split in to monthly or weekly sequences for faster processing, with the precision dependent on the quantity and detail of input time step data. The software brings possibilities for integration of multiple production sites, thus it can be used for large scale system planning, district heating, cooling and so on.

CYCLE TEMPO [21] is a program for thermodynamic modeling and optimization of systems for the production of electricity, heat and refrigeration, developed at Delft University of Technology (TU Delft). It is used for stationary operating conditions of energy and refrigeration systems. It has a combination of sequential and simultaneous solver [4,5]. The primary aim of Cycle-Tempo is to calculate the size of the relevant mass and energy flows in the system. It is possible to set the number of types of apparatuses, and the way in which they are interconnected. The program contains of a large number of models of apparatuses and pipes for creating a system model. This almost unlimited flexibility is a significant advantage over many existing programs in which the system configuration cannot be varied, or only to a limited extent. Cycle tempo input dialogs are used to enter pressures, temperatures and compositions of working fluids. Calculated energy and exergy values are used for assessment of the properties of the modeled system. If a calculation is successful, you can interactively create Q, T and value diagrams of one or more heat-exchanging apparatuses, as well as state diagrams (Mollier and T, s diagrams) of (part of) the process.

DNA [22] is a energy system simulation tool developed at the Technical University of Denmark, available free of charge, capable of handling models of any kind of energy system based on the control volume approach [2, 23]. It can handle both steady state and dynamic simulations as well as optimization of several types of energy systems: steam turbines, gas turbines, fuel cells, gasification, refrigeration and heat pumps for both fossil fuels and different biomass types. It has a sparse matrix based simultaneous solver [2,4] and is capable of handling discontinuities in dynamic equations. The component library consists of Heat

exchangers, burners, gasifiers, turbo machinery, fuel cells, fuel pretreatment, energy stores, engines valves and controllers. Additional components may be implemented written in Fortran 77. It does not feature a graphical user interface, on the contrary, it can be referred to as a modeling language or as a command line program which passes and compiles the language. It is well documented and available online free of charge.

ENERGYPLUS EnergyPlus is building performance simulation program that has its roots in the BLAST and DOE–2 programs. It was originaly written in Fortran 90 BLAST (Building Loads Analysis and System Thermodynamics) and DOE–2 were both developed and released in the late 1970s and early 1980s as energy and load simulation tools. EnergyPlus is an energy analysis and thermal load simulation program, developed by the US Department for Energy. It is very well documented. The EnergyPlus program is a collection of many program modules for calculating the energy required for heating and cooling a building using a variety of systems and energy sources by simulating the building and associated energy systems when they are exposed to different environmental and operating conditions. The core of the simulation is a model of the building that is based on fundamental heat balance principles. EnergyPlus is an integrated simulation tool which means that all three of the major parts, building, system, and plant, must be solved simultaneously, unlike its predecesors DOE-2 and BLAST which had sequential solvers. The simulation can be based on user defined sub hourly time steps. There are many modules which can be included in the simulation. The program requires google sketch up with an Open studio plug in for building modeling and zoning. It has an open code, which makes it flexible for user created modules and components, and the provided IDF editor can be used to review and edit existing modules. It has modules for simulationg: surface heat balance; climate and weather, sky, solar/shading; solar radiation reflected from external surfaces, daylighting and window calculations, air heat balance, building system simulation system; loop based HVAC, equipment sizing etc; demand limiting, on-site generation; Atmospheric pollution calculations that predict CO2, SOx, NOx, CO, particulate matter, and hydrocarbon production for both on site and remote energy conversion and some economics, however it cannot perform LCA analysis, however the output data can be fed to another program that does [24-26].

EES (Engineering Equation Solver) [2,4, 27] is a general equation solver. It gives the numerical solution of a set of algebraic equations, differential and integral equations, it can be used for optimization, provide uncertainty analyses and linear and non-linear regression and generate publication-quality plots. IT also has many built-in mathematical and thermo-physical property functions useful for engineering calculations. It is mainly used for steady state simulation, and uses a robust implementation of the Newton matrix techniques. Although it can be used for dynamic simulations as well, the implemented ODE solver has limited efficiency.

MATHLAB/SIMULINK [2,4, 28] is a commercial tool for modeling, simulating, and analyzing dynamic systems. Simulink provides a graphical user interface (GUI) for building models as block diagrams Simulink includes a comprehensive block library of sinks, sources, linear and nonlinear components, and connectors, but it is also possible to customize and/or create new blocks. It is mainly indented for control system design and not well suited for closed loop systems, which is often the case with energy systems[2]. It has access to mathematical features available in Mathlab.

Aspen HYSYS [2,4,29] process modeling tool for conceptual design, optimization, business planning, asset management, and performance monitoring for oil & gas production, gas processing, petroleum refining, and air separation industries. offers a comprehensive

thermodynamics foundation for accurate calculation of physical properties, transport properties, and phase behavior for the oil & gas and refining industries, and a Comprehensive library of unit operation models including distillation, reactors, heat transfer operations, rotating equipment, controllers, and logical operations in both the steady state and dynamic environments, GHG calculations. It has modular operation combined with the sequential solution algorithm.

WINDALI is a general modeling and simulation system capable of solving a system of semi-explicit differential algebraic equations [2,4], initial value problems, and can handle discontinuities. It has two sub-programs: Free Pascal Editor and Simulation. It uses dll (dynamic link libraries). It is available free of charge online. It is very flexible since Models can be created in practical any programming language (Pascal, C++, Fortran...), and it is possible to a user-supplied solver instead of the accompanying equation solver. It was developed at the Technical University of Denmark and is available online free of charge [30]

GATE CYCLE [31,32] is a commercial software developed by "GE Enter" software company for analyzing design and off-design performance of combined cycle plants, fossil boiler plants, nuclear power plants, cogeneration systems, combined heat-and-power plants, advanced gas turbine cycles and many other energy systems. It can be used for quick assessments, detailed engineering, design, retrofitting, re-powering and acceptance testing. Its has a modular "component-by-component approach" and advanced macro which make it flexible for modelling different system configurations. It can handle the following problems: Conceptual plant design, detailed design work, Parametric studies, Feasibility studies, Pro-formas, "What-if" analysis, Plant retrofits/re-powering. It can also be used for financial optimization of thermal power plants. It can be used for steady state simulation which is carried out via sequential solver [4,5].

UNITED CYCLE software was developed by the "Power Machines" (PM) corporation and the Technical University in Sankt Petersburg for the PM steam and gas turbine manufacturer [4]. The software is based on more than two decades of experience in computer simulations, developed component models and algorithms, and accumulated experimental data. It provides accurate calculations of steady state regimes for detailed energy flow paths of a plant for a complete range of operational loads. A sequential modular algorithm solves the mass and energy balances for both the overall system as well as all individual components, and detailed reports for each are generated automatically. The software was designed for investigating analyzing and designing of real thermal power plants, with much detail regarding fluid flows in all operating load regimes. Guarantee characteristics and power curves as well as technical boundaries of the working regime parameters are calculated using United Cycle. Since the manufacturer has typical structural configurations, the content of heat flow schemes and a set of equipment with known characteristics, the optimization of technical schemes is carried out with a goal of finding the best solution out of a few technically reasonable configurations [4,5].

Invert simulation tool is a "dynamic bottom-up simulation tool" applicable on existing building stock. It can simulate heating, cooling, domestic hot water, solar thermal systems, rational use of energy as well as renewable energy sources (RES) based on electricity supply, heat production (RES-CHP) and bio fuel production. It is designed for comparative and quantitative sensitivity analysis of promotion schemes for rational energy use, and renewable energy utilization, as well as GHG reduction. The program models the decision making process for the investor, taking into account market restrictions, technical and non technical barriers, incentives etc via "dynamic resource curve approach" [35].

HOMER (Hybrid Optimization software tools of hybrid systems) is a software developed by NREL (National Renewable Energy laboratory, USA) [33]. It is the most used optimization software for hybrid systems. It is capable of optimization of hybrid systems consisting of photovoltaics, batteries, wind turbines, hydraulic turbines, Ac generators, fuel cells, electrolyzers, hydrogen tanks, AC-DC bidirectional converters and boilers. The loads can be AC/DC ,hydrogen loads and thermal loads. The simulation is based on one hour intervals during which the input parameters remain constant. HOMER models a power system's physical behavior and its life-cycle cost, which is the total cost of installing and operating the system over its life span. HOMER can model grid-connected and off-grid systems serving electric and thermal loads, and comprising any combination of photovoltaic (PV) modules, wind turbines, small hydro, biomass power, reciprocating engine generators, microturbines, fuel cells, batteries, and hydrogen storage. HOMER performs three principal tasks: simulation, optimization, and sensitivity analysis. It does not model electrical transients or other dynamic effects, which would require much smaller time steps instead, HOMER models primary load (electric demand that must be served according to a particular schedule), deferrable load (electric demand that can be served at any time within a certain time span) and thermal load (demand for heat) The fundamental principle that HOMER follows when dispatching the system is the minimization of cost. HOMER represents the economics of each dispatchable energy source by a fixed cost per hour, and a marginal cost of energy per kW/h, which represent all costs associated with producing energy with that power source during that hour [33].

4. Conclusion

Many decisions about the configuration of the system need to be made during energy system design: What components does it make sense to include in the system design? How many and what size of each component should you use? What operating strategy is the most efficient? What operating strategy offers the greatest cost saving?

There is a large number of technology options and the variation in technology costs and availability of energy resources make these decisions difficult. Many of the reviewed software try to help in the mentioned modeling phase. There is a number of software solutions as tools for determining energy performance of a system or a building (Object) and help in decision making. After the decision making phase, based on the chosen technologies for integration, many software give possibilities for system performance optimization in different operating regimes and/or evaluation based on LCA, costs, energy savings, GHG emissions and other economic and financial factors.

Most of the reviewed programs have to a graphical user interface, however, some like Windali and Insel for example, have command line interface which may be more difficult for the end user. Some of them have sequential solvers (BLAST, DOE-2, Aspen Hysis, Gate Cycle, Camel), simultaneous solver (EnergyPlus), or the combination of sequential and simultaneous solver (Cycle tempo, Energy Pro). There are also general solvers, which can simulate any model given by a set of equations, such as Insel, Windali, EES.

Many of the software reviewed focus on simulating, estimating and optimizing energy system coupled with a building (object), thus analyzing its performance, such as Blast, BSim, DOE-2, Trnsys, Energy Pro, EneegyPlus, and A building performance thermal analysis tool by Mendes et. al. There is number of programs for simulating hybrid systems (Hybrid2, Homer, Trnsys, and some of them are more oriented towards to solar technologies,

for example Insel, HOMER. Programs suitable for simulating power plants are DNA, GATE CYCLE, CogenPro.

Most of the software are designed for steady state simulations, however many use a quasi dynamic approach where the variables have constant value during a given time step (DOE-2, Energy Pro, HOMER, HYBRID, HOGA, CAMEL) which is usually set to one hour or can be user defined in some cases. Software capable of dynamic simulations are DNA, TRNSYS, Simulink, DNA, Bsim. Nevertheless, for many processes, the quasi static time steps are sufficient for engineering purposes due to the physical nature and inertia of the process.

Most of the programs come with a modular structure, which makes them flexible and suitable for a variety of system configurations. Some, like Ecotect, gproms, insel, Trnsys, Hybrid2 or open code for adding user made components to the system configuration like Trnsys, DNA, Energy plus, Mathlab/Simulink, WinDali, making them suitable for simulation of almost any system configuration.

Software capable of doing optimization is TRNSYS, ENERGY PRO, DNA, ENERGY PLUS, EES, Aspen HYSIS, Homer. Some of the software take economic and/or financial factors for the optimization, for example HYBRID2, Invert, Aspen HYSIS, EnergyPRO, RetScreen, Doe-2. The programs that is capable of performing LCA analysis is HOMER.

Generally, energy modeling is a solid tool for evaluation relative energy use, and not for predicting absolute energy use, since an estimation can be made on the performance various types of technologies, whilst an actual estimation energy performance cannot be made because this is usually beyond the predictive capabilities of the software since it is influenced by many uncertain factors.

References

1. Elmegaard B., N. Houbak, *Software for the Simulation of Power Plant Processes - Part A*, ECOS 2002, Berlin, 2002.
2. Elmegaard B., N. Houbak, *Software for the Simulation of Power Plant Processes - Part B*, ECOS 2002, Berlin, 2002.
3. Elmegaard B., N. Houbak, *Robust Implementation of Process Simulators and Their Associated Models*, ECOS 2000, Enschede, The Netherlands, 2000.
4. Zivkovic D., Software for the simulation processes in CHP plants.
5. Zivkovic D., Simulation of the cogeneration plant "Avedoreverket" – Unit 1 with DNA, Gate Cycle and United Cycle programmes.
6. Crawleya B., J. W. Handb, Michae¨ l Kummertc, Brent T. Griffithd , Contrasting the capabilities of building energy performance simulation programs, Building and Environment 43 (2008) 661–673.
7. Christensen C., R. Anderson , S. Horowitz, A. Courtney, and J. Spencer , BEopt™ Software for Building Energy Optimization: Features and Capabilities, Technical Report, NREL/TP-550-39929 August 2006, Prepared under Task No. BET6.8004
8. Pedersen C.O., D.E. Fisher, R.J. Liesen, R.K. Strand and R.D. Taylor, ENERGYBASE: THE MERGER OF BLAST AND DOE–2
9. http://www.psenterprise.com/gproms/index.html , 03.10.2011.
10. Peter Tijl , Assessment of the parameter estimation capabilities of gPROMS and Aspen Custom Modeler, using the Sec-Butyl-Alcohol stripper kinetics case study, Graduation report, Eindhoven Technical University, Department Mechanical Engineering, Dynamics and Control Technology Group, Amsterdam, July 2005; DCT 2005.96

11. http://www.retscreen.net/ang/home.php ; 15.10.2011. (clean energy project analysis: retscreen® engineering & cases textbook, introduction to clean energy project analysis chapter, Minister of Natural Resources Canada 2001-2005. ISBN: 0-662-39191-8)
12. http://www.insel.eu; 13.07.2011.
13. Eicker U., D. Pietruschka, J. Schumacher, J. C. Fernandes, T. Feldmann, E. Bollin, Improving the energy yield of PV power plants through Internet based simulation, monitoring and visualisation. Proceedings 20th European Photovoltaic Solar Energy Conference (2670-2673) Barcelona, Spain, June 2005.
14. http://usa.autodesk.com, 15.08.2011.
15. http://www.trnsys.com, 15.05.2011.
16. Matthew J., Duffy, Marion Hiller, David E. Bradley, Werner Keilholz, Jeff W. Thornton1 TRNSYS – FEATURES AND FUNCTIONALITITY FOR BUILDING SIMULATION, Eleventh International IBPSA Conference Glasgow, Scotland July 27-30, 2009.
17. Hiller M., S. Holst, A.Knirsch, M. Schuler, Trnflow, a new tool for the modelling of heat, air and pollutant transport in buildings within trnsys, Eighth International IBPSA Conference Eindhoven, Netherlands, August 11-14, 2003.
18. Wetter M., C. Haugstetter, Modelica versus trnsys – a comparison between an equation-based and a procedural modeling language for building energy simulation, Second National IBPSA-USA Conference Cambridge, MA, August 2-4, 2006.
19. Breckman W., L. Broman, A. Fiksel, S. Klein, E Lindberg, M. Schuler and J. Thornton, Trnsys the most complete solar energy system modeling and simulation software, Renewable Energy, Vol 5. Part I, pp 486-488, 1994.
20. Lira L., Hybrid heating and cooling system optimization with TRNSYS, MSc in Energy Systems and the Environment theses, University of Strathclyde Engineering, September 2008.
21. http://www.emd.dk/energyPRO/Frontpage
22. http://www.et.web.mek.dtu.dk/software/DNA/index.html
23. Elmegaard B., N. Houbak, B. Lorentzen, DNA - a Network-based Energy System Simulator, FLOWERS'97, Florence, Italy, 1997.
24. http://apps1.eere.energy.gov/buildings/energyplus/
25. Crawley D.B., L.K. Lawrie, C.O. Pedersen, R.J. Liesen, D.E. Fisher, R.K. Strand, R.D.
26. Taylor R.C., Winkelmann, W.F. Buhl, Y.J. Huang, A.E. Erdem. ENERGYPLUS, A.
27. New-Generation Building Energy Simulation Program. Proceedings of Building Simulation '99, Volume 1: 81-88.
28. Drury B. C., L. K. Lawrie, C. O. Pedersen, F. C. Winkelmann, EnergyPlus: Energy Simulation Program, ASHRAE journal online, April, 2000, vol. 42, no. 4, pg. 49-56.
29. http://www.mhhe.com/engcs/mech/ees/na.html; 15.08.2011.
30. http://www.mathworks.com/products/simulink/
31. http://www.aspentech.com/products/aspen-hysys.aspx
32. http://www.et.web.mek.dtu.dk/WinDali/Index.html
33. http://www.ge-mcs.com/en/bently-nevada-software/performance/gate-cycle.html
34. Sergei N. Romanov, Anatoly G. Kutakhov and Konstantin S. Romanov, "Software united cycle for simulation of static operation modes of power plants", Proc. SPIE 4627, 306 (2002); doi:10.1117/12.456288
35. http://www.homerenergy.com/software.html

Inverse Calculations In Prevent Crisis Phenomenon of Socio-Economic Systems

I.V. Borshchuk, B.E. Odintsov, V.D. Shkvir

Abstract. Possibilities of the inverse calculations apparatus in management of economic systems stability were shown. The apparatus of inverse calculations was demonstrated on the example of the impact of some bifurcation point system.

Keywords: inverse calculation, stable development of socio-economic systems, point of bifurcation system, attractors, dissipative processes.

1. Introduction

Classical science of forecasting, based on a causal paradigm did not justify the expectations laid on her in most cases, because great number of shock humanity crises, breaking out in XX and beginning of XX1 of centuries, were not foreseen. Different on a scope and depth (global and local, financial and technogenic, political and social, ethnic and intergovernmental), they generated disbelief in a scientific prediction in a long-term prospect.

Pessimism gradually began to be transformed in a hope and confidence as far as the transfer of another look to the evolution of the socio-economic systems earlier known under the name Physics of non-equilibrium processes, the establishment of which led to the emergence of concepts of self-organization and self-development systems. Later, due to H. Haken, this scientific direction got the name of synergetics [2], the main task of which consists of search steady, adequate methods of organization the processes in system. Such approach is new for an economy and sociology but has been successfully used in mathematics, physics, chemistry, cosmology, ecology, biology, etc. However in an economy this approach carries especially theoretical character while, remaining at the level of quality recommendations. The real report carries practical character in which synergetics ideas will be realized by reverse calculations.

2. Formulation of the problem

A new look at the development of socio-economic systems requires their consideration through the prism of undeniable fact that all natural laws are reduced to the corresponding laws of maintenance. It is possible to do an important conclusion: any behavior of the system is sent to self-preservation, i.e. on that stable state which will allow her further existence and development. The search of the stable state is possible if the reflection behavior of the system using nonlinear dynamic models, usually containing various kinds of nonlinear differential equalizations.

Study and analysis of basic ideas of this scientific direction are theories of bifurcations, theory of catastrophes, theory of complication, fractals and chaos shows that the main efforts of scientists focused on prediction of the development process or the behavior of systems. However to predict the future state of an unstable system and subsequently uncontrolled change in the trajectory of its development, is not enough. Timely interference is required to prevent the crisis. It should be emphasized – this is not about maintaining the myth, as external affecting on difficult system the trajectory of her behavior can be changed in accordance with the wishes of subject without every account of her nature (structures). Times of such illusions had passed. Practice showed that influencing on the system of strict

measures can be achieved only short-term effect, because the system sooner or later will come to one of the stable states, which adequate to her nature. Therefore we must know the nature of the system its structure that will allow with minimum efforts direct the process of her development exactly to her. Intervention should not be ″administrative″, but ″soft″, that is taking into account a structure which provides the stable state of the system. In our view not enough attention is spared to this problem that requires the search of ways to fill this gap.

3. Results of researches

Dissemination of ideas and principles of simulation processes of evolution and self-organization of biological, chemical and other systems using nonlinear dynamics on the socio-economic phenomena required for analyze the number of concepts before in an economy not used. The points of bifurcation, attractors and dissipative processes belong to such concepts.

Under bifurcation of some system is understood transformation of her phase portrait as a result of changes in parameter values. The point of bifurcation determine by overcoming the parameter of his threshold in result what system will be in the unstable state, determines. The same actions can cause opposite results even at small changes in critical, it is the bifurcational area of system parameters. In the point of bifurcation the system qualitatively changes its behavior heading to one of regular attractors.

In basis of concept the attractor is a fundamental aspiration of all the systems to self preservation. All pure exists so far as it is able to resist external influences in the gradual moving toward the stable state. Therefore under the attractor in generally understood that stable state which is currently aspiring system. As it applies to the socio-economic systems it is possible to use the aim formulated subjectively.

Overcoming of bifurcational point can come natural by the not controlled way (by the method of tests and errors) or purposefully on basis in time generated management abilities of the for adjustment the trajectory of system behavior. Not controlled process is fraught with the system to reach an erroneous path that is one of the many unsustainable development trajectories that will eventually be destroyed by dissipative processes. Dissipative processes that were imposed by I. Prigogine are chaotic processes leading to dissipation of energy, matter and information [1]. They serve as the instrument of ″natural selection″ correct path of development, which corresponds to one of the attractors.

This raises justifiable desire to accelerate the system's output to the desired attractor, saving time and costs for development of unstable structures and their subsequent destruction of the dissipative processes. For this purpose in the point of bifurcation or its outskirts is necessary to carry out fully certain changes in the values of factors because of which depends subsequent development of the system avoiding in the same time the phase of tests and errors.

Directing a process in a necessary river-bed is possible by means of reverse calculations producing managing measures in determined and stochastic environments. Theory of such calculations was expounded in [3], applicable also and for forming of decisions in an indefinite environment.

The task of the bifurcation theory, built on the basis of a qualitative analysis of singular points of a dynamical system is to identify conditions for stability of the system, depending on the parameters of the system. Some values of last, named bifurcational, entail quality changes in behavior of the system. In the vicinity of the bifurcation point, system after some stay in a state of instability continues the motion in the direction of the attractor

chosen by chance. To make this choice was more directed, we define the behavior of the system, depending, for example, two variables, using nonlinear equations of the form

$$F(x, y, u, u_x, u_y) = 0, \quad (1)$$

where x, y - the variables defined in some space;

$u(x, y)$ - unknown function;

u_x, u_y - partial derivatives on x, y;

If a task solves analytically, then it is possible to get a decision as a general integral that is equal

$$\Phi(x, y, u, r) = 0 \quad (2)$$

where r - a parameter.

Further, as a result of qualitative analysis of this equation is defined threshold value of parameter r which is a bifurcation point. Through this point can pass two or more branches of decision of this equalization. It is known that at passing of parameter through a bifurcational value can occur splitting for a few special points or may be a new cycle [4].

For the search of ″soft″ managing actions on the behavior of the object should first enter the initial (actual) values of variables and determine the value function in some approximation to the bifurcation point:

$$u_0 = u(x_0, y_0, r_0 - \varepsilon), \quad (3)$$

where $r_0 - \varepsilon$ - the bifurcation parameter value in some his vinicity.

To obtain the control actions for the formulation of the problem of inverse calculations additional information of subjective property is needed, namely: increase of function $\pm \Delta u$; reflecting the desired approximation to the attractor, the priorities in increments of variables is equal α and β, as well as trends in the changes of these variables (plus or minus). Then the gotten decision will be complemented by the following information:

$$u_0^{\pm} = u(x_0^{\pm}(\alpha), y_0^{\pm}(\beta), r_0 - \varepsilon), \quad (4)$$

where $x_0^{\pm}(\alpha), y_0^{\pm}(\beta)$ - positive or negative increase of variable depending on the prioritization of the ways to achieve goals.

To find the control actions necessary to solve the following problem of the inverse calculation:

$$\begin{cases} u_0 \pm \Delta u = u(x_0 \pm \Delta x, y_0 \pm \Delta y, r_0) \\ \dfrac{\Delta x_0}{\Delta y_0} = \dfrac{\alpha}{\beta} \end{cases} \quad (5)$$

where $\pm \Delta u$ - purpose of controlling the behavior of the system (desired gain of the function $u(x, y)$);

$\pm \Delta x, \pm \Delta y$ - the unknown increments of variables having maintenance of managing influences.

The example with the prepared decision of differential equalization consider in [5].

The brought raising over of tasks require knowledge of the responses to a number of important issues, for example: How to identify the future structure of the system that best suits its nature? How to estimate and choose a socio-economic attractor that by the best character corresponds to the guided process? What is necessary to undertake, if passing the parameter through a bifurcational value breaking up of the special point on a few happened? How to set the problem of inverse calculations, if you have multiple parameters? Answers for these questions will allow develop the real levers of influence on the processes of socio-economic nature.

4. Conclusions

Let's note that always person will need to influence the course of certain events, processes or objects. And if these processes or events can be presented analytically (differential or by other equalizations) and the aims of influence are known then the vehicle of reverse calculations that shown on the example of influence in some bifurcational point, will be claimed in any case.

References

1. Пригожин И. Конец определенности. Время, хаос и новые законы природы / И.А. Пригожин. - Ижевск: НИЦ «Регулярная и хаотичная динамика», 2001. – 216 с.
2. Хакен Г. Принципы работы головного мозга: синергетический подход к активности мозга, поведению и когнитивной деятельности / Г. Хакен. - М., 2001. – 353 с.
3. Одинцов Б.Е. Обратные вычисления в формировании экономических решений / Б.Е. Одинцов. - М.: Финансы и статистика, 2004. – 192 с.
4. Амелькин В.В. Дифференциальные уравнения в приложениях / В.В. Амелькин. - М.: Едиториал УРСС, 2003. – 208 с.
5. Б.Е. Одинцов. Обратные вычисления как средство обеспечения устойчивого развития социально-экономических систем / Б.Е. Одинцов, В.Д. Шквир // Науковий вісник національного лісотехнічного університету України: збірник науково-технічних праць. – Львів: РВВ НЛТУ України. – 2010. Вип. 20.7. – с. 254 – 260.

Information Technologies and Student Practices

Rosen Kirilov[1]

[1] UNWE, 1700 Sofia, Bulgaria
rosenkirilov@mail.bg

Abstract. The project BG051PO001 - 3.3.07-002 "Student's practices" implemented by the University of National and World Economy in partnership with the Ministry of Education and is co-financed by the European Social Fund. Its objectives include improving the quality level of education by providing opportunities to gain experience and improve the practical skills of students of University of National and World Economy in accordance with the needs of the labor market. The paper examines aspects of information management and monitoring of the project and formulates conclusions.

Keywords. Information technologies, web technologies, student practices.

1. Introduction

Implementation of activities under the project requires active involvement of the various information and communication technologies. One of the major information problems of the project is the option for students who have been approved by employers to generate their contracts without fulfilling some of the requirements of the Ministry. Similar approach to freedom when signing contracts transmitted flexibility of the project, but without excessive effort by the university staff to monitor individual cases. This has a direct impact on verification costs. The paper addresses the issues of prior approval of the student contracts with specially developed software.

2. Approach to pre-approval of contracts

Approach is based on the following main actions [2] (Figure 1):

- Sending a student request;
- Verification of the request;
- Verification of student status (Figure 2).

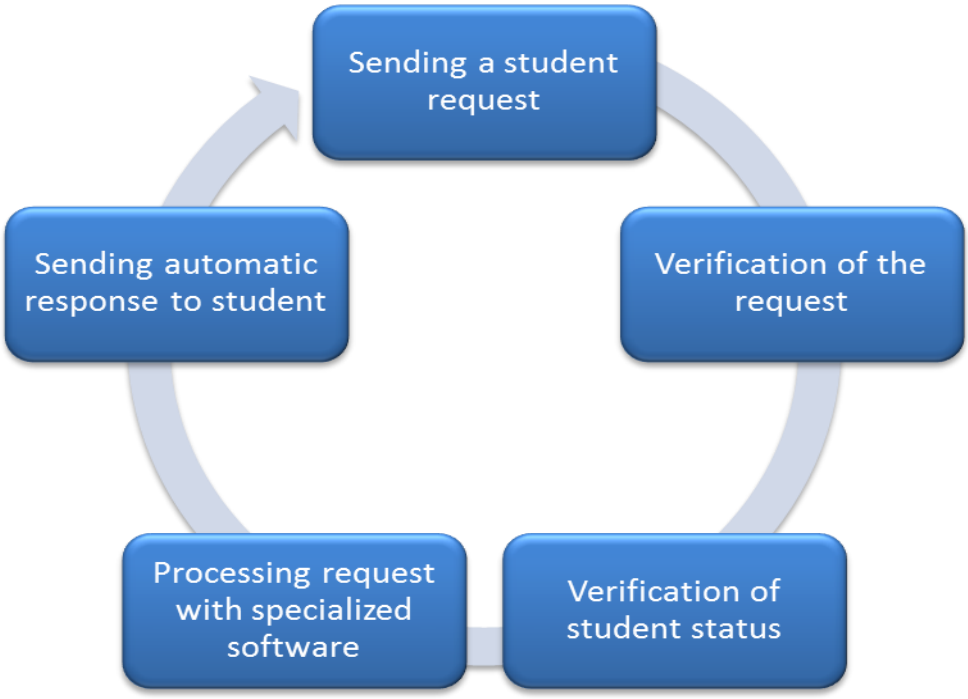


Figure 1. Pre-approval of contracts

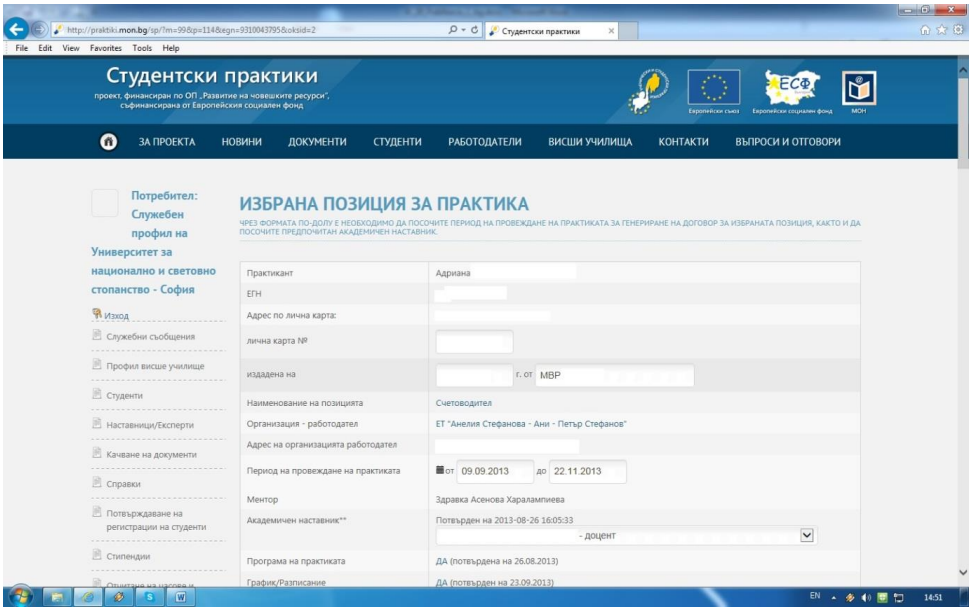


Figure 2. Student status

- Processing request with specialized software;

- Sending automatic response to student.

3. Software for pre-approval of contracts

For pre-approval of contracts University of National and World Economy has developed specialized software. For its development are attached some of the technological specifics described by Milev [1]. At Figure 3 is presented the design of the main screen of this software.

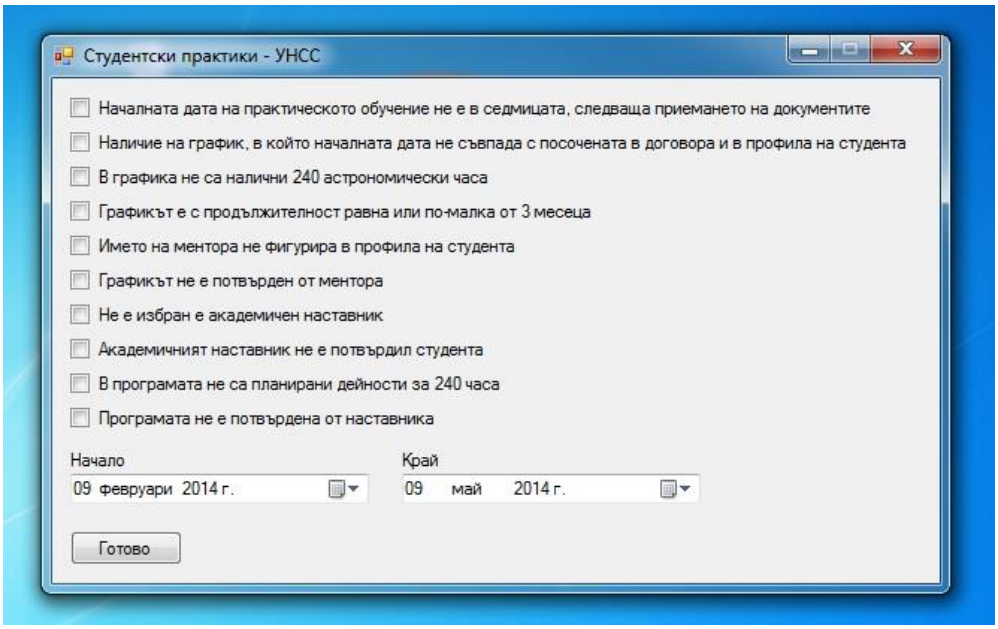


Figure 3. Main screen of the software

With the use of this software an automatic response is generated that is sent to the e-mail address of the student. Below is presented a fragment of the source code of the software.

```
[STAThread]
        private void CopyToClipboard()
        {
            bool checkOK = true;
            int i = 1;

            foreach (Control c in this.Controls)
            {
                if (c is CheckBox)
                {
                    CheckBox cb = c as CheckBox;

                    if (cb.Checked == true)
                    {
                        content = content + i + "." + cb.Tag + Environment.NewLine;
                        checkOK = false;
```

```
                        i++;
                    }
                }
            }

            string finalText = content + Environment.NewLine;

            if (checkOK)
            {
                finalText = this.home + Environment.NewLine + contentOK1 +
this.signature + Environment.NewLine + contentOK2;
            }
            else
            {
                finalText = this.home + finalText + this.endNotOK +
Environment.NewLine + this.signature;
            }

            Clipboard.SetText(finalText);

            foreach (Control c in this.Controls)
            {
                if (c is CheckBox)
                {
                    CheckBox cb = c as CheckBox;

                    cb.Checked = false;
                }
            }

            this.dateTimePickerBegin.Value = DateTime.Today;
            this.dateTimePickerEnd.Value = DateTime.Now.AddMonths(3);
        }
```

4. Conclusion

The main purpose of University of National and World Economy with this project is to provide students with knowledge, skills and competencies as a result of training and work in a real practical environment that creates opportunities for better realization in accordance with the requirements of the labor market. For the period of the project University of National and World Economy provides to conclude the contract with about 3,000 students.

References

1. Milev, P., "Technological issues of implementing software for media monitoring", "Trakia Journal of Sciences", Vol. 11, Suppl. 1, 2013, Trakia University, ISSN 1313-7069, 2013.
2. Official website of project "Student practices", http://praktiki.mon.bg/sp, 02.12.2013.

Problems with Electronic Management of Public Projects

Rosen Kirilov[1]

[1] UNWE, 1700 Sofia, Bulgaria
rosenkirilov@mail.bg

Abstract. The project BG051PO001 - 3.3.07-002 "Student's practices" implemented by the University of National and World Economy in partnership with the Ministry of Education and is co-financed by the European Social Fund. Its objectives include improving the quality level of education by providing opportunities to gain experience and improve the practical skills of students of University of National and World Economy in accordance with the needs of the labor market. The paper examines aspects of information management and monitoring of the project and formulates conclusions.

Keywords. Information technologies, public projects, electronic management

1. Introduction

The project BG051PO001 - 3.3.07-002 "Student Practices" (the project ends on October 31, 2014) is implemented by the University of National and World Economy in partnership with the Ministry of Education and is co-financed by the European Social Fund [3]. Its objectives include improving the quality level of education by providing opportunities to gain experience and improve the practical skills of students of University of National and World Economy in accordance with the needs of the labor market.

2. Process of contracting

Applications, conducting and reporting of the project is fully built web-based platform. In this way, the project creates a good practice for e-government public works project, which is a factor for effective administrative policy similar to that described by Kirilova [1]. For the period from July 1, 2013 to February 10, 2014, the number of contracts for student practice in the University amounted to 2110 units (Table 1):

Table 1. Number of contracts

Date	Registered students within the information system of the project	Number of practices confirmed by employers	Number of practices confirmed by students	Number of contracts
04 October 2013	4073	1479	1188	770
31 January 2014	6163	3136	2969	1930
07 February 2014	6408	3262	3101	2110

A graphical representation of the registered student is given in Figure 1.

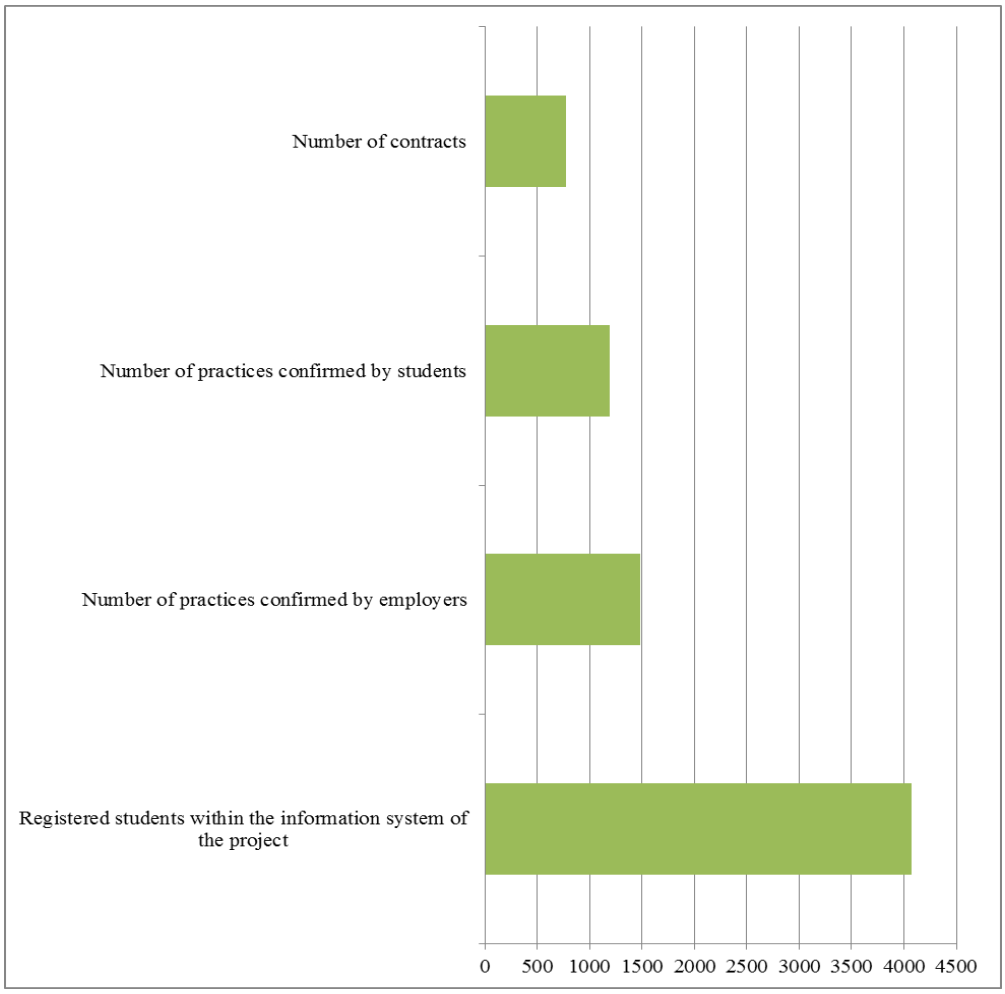


Fig.1. Number of registered students and signed contracts

3. Trends and forecasts

Table 2 shows the distribution of the number of contracts awarded by weeks of project start 07.08.2013 until 13.01.2014 inclusive.

Table 2. Number of students started their practical training by weeks

Date	Number of started
8.7.2013	47
15.7.2013	49
22.7.2013	33
29.7.2013	61
5.8.2013	32
12.8.2013	46
19.8.2013	34
26.8.2013	29
2.9.2013	42
9.9.2013	18
16.9.2013	19
23.9.2013	18
30.9.2013	56
7.10.2013	61
14.10.2013	65
21.10.2013	114
28.10.2013	114
4.11.2013	95
11.11.2013	69
18.11.2013	78
25.11.2013	86
2.12.2013	138
9.12.2013	78
16.12.2013	92
23.12.2013	15
30.12.2013	40
6.1.2014	5
13.1.2014	109
Total	**1643**

To reflect the estimate for completion of the student to September 30, 2014 in the formation of the forecast provided contracting to July 31, 2014 Figure 2 shows the graphical distribution of the number of contracts awarded to students by the weeks tend to July 31, 2014.

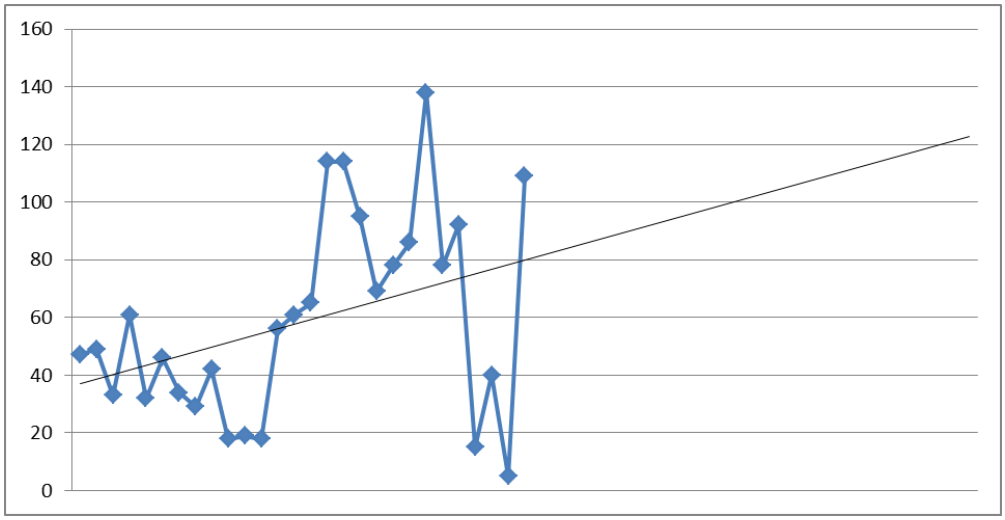


Figure 2. Numbers of contracts awarded to students by the weeks tend to July 31, 2014

Assuming thus formed is a tendency to contract it in Table 3 we can represent the expected number of students by the weeks to July 31, 2014

Table 3. Number of students who are expected to start their practical training by weeks to July 31, 2014

Date	Number of students who are expected to start their practical training by weeks
20.1.2014	101
27.1.2014	101
3.2.2014	101
10.2.2014	101
17.2.2014	101
24.2.2014	101
3.3.2014	101
10.3.2014	101
17.3.2014	101
24.3.2014	101
31.3.2014	101
7.4.2014	101
14.4.2014	101

21.4.2014	101
28.4.2014	101
5.5.2014	101
12.5.2014	101
19.5.2014	101
26.5.2014	101
2.6.2014	101
9.6.2014	102
16.6.2014	102
23.6.2014	102
30.6.2014	102
7.7.2014	102
14.7.2014	102
21.7.2014	102
28.7.2014	102
Total	**2836**

4. Conclusion

All the above gives us reason to believe that at started until now 1643 students (total contracts signed until now is 1780 units) and expected number of started to July 31, 2014 is 2836 students , the total number of completed practical training project can reach 4479 units. For the universities is an important participation in such projects and their management. It would be interesting the possibility of linking some of the data from such a system in university internal information systems, such as those for research projects. Similar approaches are described by Milev [2].An additional argument in favor of this prediction is of great interest to students of University of National and World Economy in project "Student practices" which is confirmed by the number of registered students and the number confirmed by an employer that with the onset of the summer months of 2014 can be expected to get increased.

References

1. Kirilova, K., "Importance of providing information for conducting effective local administrative policy", International Conference „SECURITATEA INFORMAŢIONALĂ 2013”, ACADEMIA DE STUDII ECONOMICE DIN MOLDOVA, ISBN 978-9975-75-640-2, 2013.
2. Milev, P., "Development of university information system for research projects", International Conference „SECURITATEA INFORMAŢIONALĂ 2013”, ACADEMIA DE STUDII ECONOMICE DIN MOLDOVA, ISBN 978-9975-75-640-2, 2013.
3. Official website of project “Student practices”, http://praktiki.mon.bg/sp, 02.12.2013.

Technological Features and Current State of Project "Student practices" at UNWE

Plamen Milev[1]

[1] UNWE, 1700 Sofia, Bulgaria
pmilev@unwe.bg

Abstract. The project "Student practices" is implemented in partnership with all 51 universities in Bulgaria. All the students included in the Register of current and discontinued students and PhD students of Ministry of Education can apply to participate in the project. Considered as successfully completed is a practical training in a real work environment within 240 hours. After completing and reporting of the student practice a scholarship is being paid. Every student has the right to be included in the practical training of the project once in their training for each degree (once as a "bachelor" and once as "master"). In the budget of the project is included funding for remuneration of academic mentors from universities and mentor employers.

Keywords. Project, student practices, UNWE.

1. Introduction

The project aims to achieve high goals, related to improving the quality of education of students, as well as their further development, shortening the distance between them and the employers. Here are the main reasons for creating this project [3]:

- The quality of education needs to be improved by providing opportunities for practical experience and improving the practical skills of University students to the needs of the labor market;
- There's no stable relationship between educational institutions and the real business, which is a premise for the inability of young people to find jobs after successfully graduating;
- Students need incentive to take the initiative to participate in extracurricular classes and to experience a real working environment;
- Universities need to modernize their curriculum in order to be most useful for the needs of the market;
- Employers lack the established mechanism by which to make qualitative selection of future business executives.

2. Basics

First of all, the candidates for participation in this project must be included in the Register of current and discontinued students and PhD students, supported by the Ministry of Education. In this regard, the work and problems of this register are similar to other public records with data reviewed by Kirilov [1]. Participation in the project can be once for each degree of the students (once as a "bachelor" and once as "master"). The practical training should mandatory match the student's specialty or professional field and be carried out within 240 astronomical hours according to a preliminary program. It should be different

from the provided in the academic curriculum of the relevant specialty or educational degree. Training is conducted by and under the supervision of an employee of the organization employer mentor and under the guidance of university professor – academic mentor. Those interested in applying can do so by registering a personal profile on the official website of the project - http://praktiki.mon.bg/sp, providing the necessary data and consistently perform the necessary steps. After successful registration, the student has the opportunity to review the automatically filtered from the web system ads that match his/her professional field and if interest in any of them he/she can apply. Follows a selection by the employing organization which includes review of the student's profile information and discretionary company chooses the candidates they prefer. Upon approval the student has 5 working days to confirm his participation in the practical training. Upon confirmation the student has to give his/her bank account, which is exclusively personally his/hers, where the scholarship will be transferred upon successful completion of training.

Upon successful completion of the project the student prepares a final report on its activities, which should be checked and confirmed by the academic mentor and the employer mentor. Electronic confirmation of the completed practice meets modern principles of methodological foundations for the development of electronic services described by Kirilova [2]. In this regard, the project is innovative and implements best practices of e-institutional management. Everyone involved in the program receives a certificate that can serve him/her in awarding academic prizes if the university has planned ones. Students that entered wrong information in the web system lose the right to participate in the project and give back the received scholarship. "Student practices" is a project that not only allows students to gain real experience in the business environment, but also helps employers to make a qualitative selection of future employees for their companies. Officers working on the project are engaged actively with information campaigns among potential employers' organizations and continuously increase their number.

3. Statistics

Let's take a look at some of the statistical distributions about the project, namely:

- Distribution of participants by gender;
- Distribution of participants by age;
- Distribution of participants by cities;
- Distribution of participants by cities outside Sofia.

Distribution of participants by gender is presented at Figure 1. In gender distribution of the participants in the project the women overweight, which is not surprising since many of them are stronger interested in education and have more responsible attitude towards it.

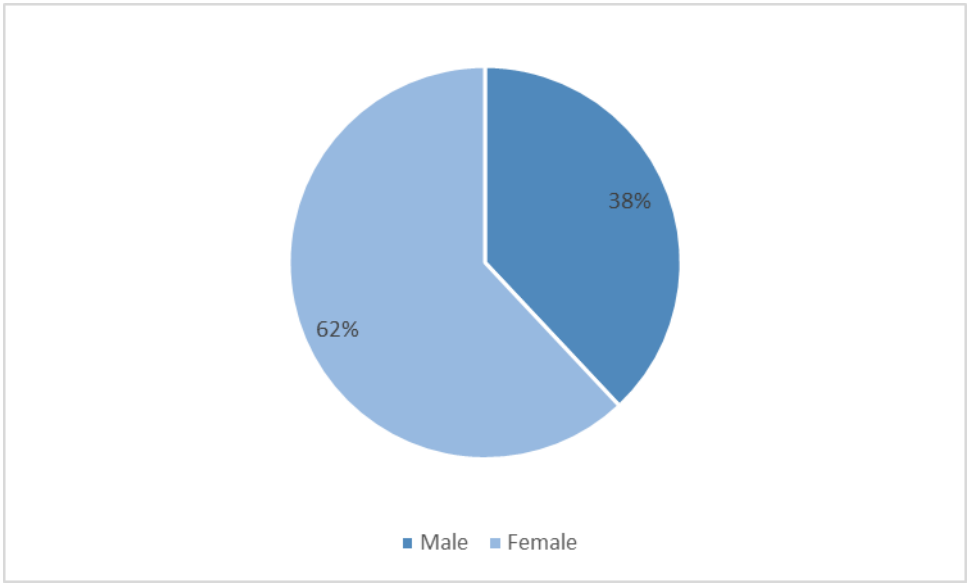


Figure 1. Distribution of participants by gender

Distribution of participants by age is presented at Figure 2.

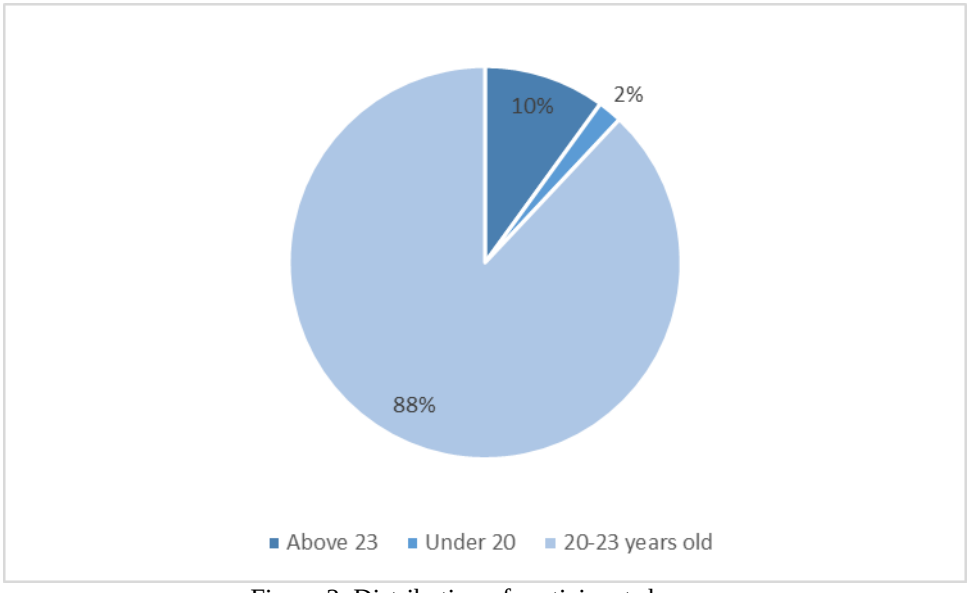


Figure 2. Distribution of participants by age

The lowest number of participants was observed in the age group with an upper limit of 20 years (only 2%). Students from this group are usually the first year of their training course and are not strongly interest in participating in provided extracurricular activities. The second group of students is between 20 and 23 years of age. They tend to show a significant increase in the interest in student practices. Students in this age range enter into their real

training in the fields have chosen and they are focused on acquiring additional skills training for the real business environment. In the last third group of students over 23, the percentage of those willing to participate in practical training is again reduced (10%). During this period, students usually are looking towards searching a real job they already have the necessary training skills for.

Distribution of participants by cities is presented at Figure 3.

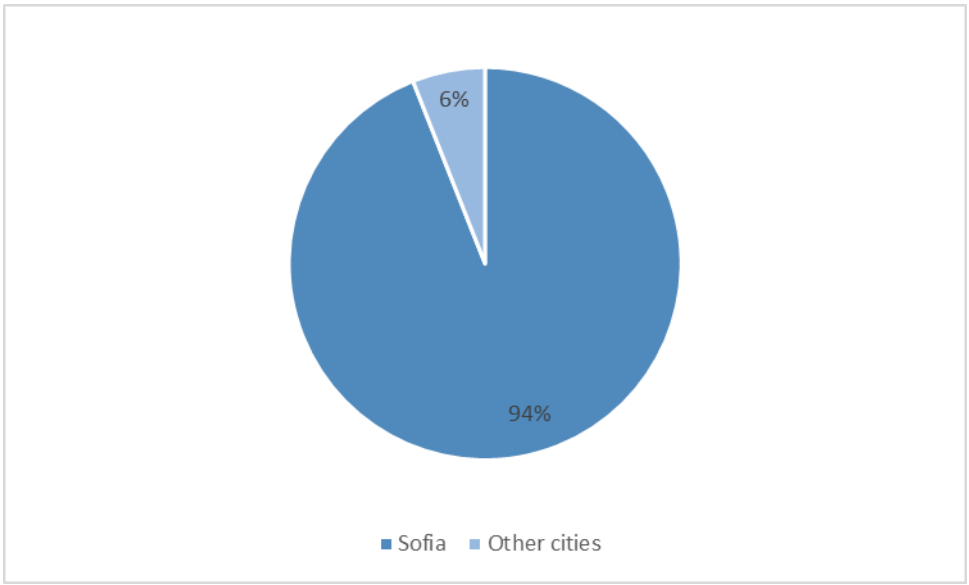


Figure 3. Distribution of participants by cities

Statistics show that 94% of students who participate in practices choose to do so in Sofia. The fact is indicative that young people prefer to be realized in the capital, believing that it offers the greatest amount of opportunities. Only 6% of participants in practical training choose a smaller city than Sofia.

Distribution of participants by cities outside Sofia is presented at Figure 4.

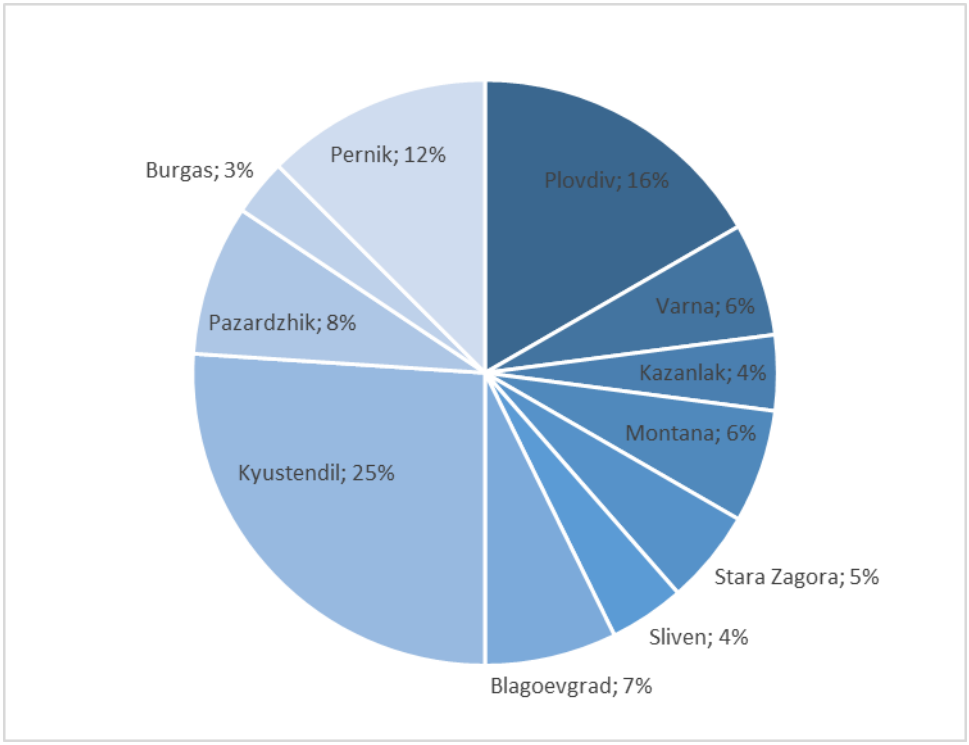


Figure 4. Distribution of participants by cities

After Sofia, as the most desirable city to conduct the practice appears Kyustendil. Surprisingly the city, which is placed at number 23 in population in Bulgaria, has a significant percentage of students studying the project, surpassing even crowded cities like Plovdiv, Varna and Burgas.

4. Conclusion

The paper examined the basics about project "Student practices" at UNWE. Statistical distributions were represented by different criteria, with data actual to 01 of November 2013. The paper confirms the opinion about the importance of this project to the public society.

References

1. Kirilov, R., "Analysis and improve the performance of certain public registers", International conference "Interaction theory – practice, key problems and solutions", Burgas Free University, ISBN 978-954-9370-80-5, 2011.
2. Kirilova, K., "Methodological issues in assessing the capacity of local authorities for developing and delivering electronic services", "Trakia Journal of Sciences", Vol. 11, Suppl. 1, 2013, Trakia University, ISSN 1313-7069, 2013.
3. Official website of project "Student practices", http://praktiki.mon.bg/sp, 02.12.2013.

Method of Storing Dynamic Data in a Relational Database on System for Scientific Research at UNWE

Plamen Milev[1]

[1] UNWE, 1700 Sofia, Bulgaria
pmilev@unwe.bg

Abstract. The paper examines the possibility of storing dynamic data in a relational database. Such a need exists in the system for scientific research at the University of National and World Economy. Under dynamic data the paper refers to data with variable structure and variable nature of different periods in terms of research projects. The paper presents a model for such a dynamic storage in a relational database and examines its advantages and disadvantages.

Keywords. Database, dynamic data, scientific research, UNWE.

1. Introduction

The university system for managing scientific research automates the actions of application, evaluation, negotiation and reporting of research projects. Its operation goes through the stages, illustrated in Figure 1. There are four different types of users in the university system for managing scientific research:

- Administrators – these are users, who should be able to modify all major nomenclatures and should be able to configure the system to work according to the budget allocation. These users are responsible for implementation of all documents related to the project during its life cycle. It is possible for the administrators to obtain information, monitor and inform other users about expiring terms on certain tasks where they are overdue. Administrators have rights for preparation of reports and statements.
- Project owners – users of this role have the rights to submit project proposals and for each stage of the project to introduce and edit data in the project. All the adjustments and corrections are only possible within the university regulations.
- Reviewers – these users have access to certain projects and possess the possibility to give reviews, including text reasoning, numerical estimates and conclusions on the projects.
- Members of the research board – users of this role have access to all the records in project proposals and also to the reviews of the projects. Their rights include rating and ranking of the projects, negotiation and acceptance of reports.

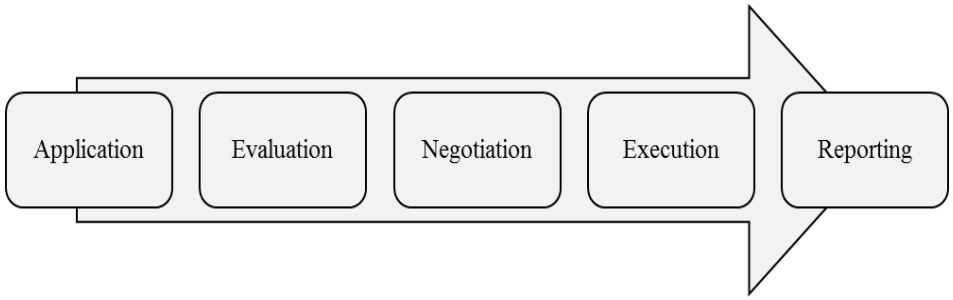


Figure 1. Stages of operation of the university system for managing scientific research

These stages are similar to those, described by Kirilova in terms of project management in local government [1]. At each of these stages the information system works with different data. The aim of this paper is to examine the way the data for research projects is stored within the database. Specific for the data is that it has a different structure for each calendar year. For this reason, a relational database should often change its structure in order to determine proper functioning of the system. The paper presents a method of storing data on the principle of attribute – value, where the software part cares for the integrity of the data and the database takes care only for the storage.

2. Relational part

A portion of the database is for a description of standard nomenclatures. These include:

- Users – comprise four above described roles in the system. Users are characterized by degree, academic position, department, full name and email.
- Degrees – each user can have a degree. Degree is characterized by name and abbreviation.
- Academic positions – each user can take an academic position. Academic position is characterized by name and abbreviation.
- Departments – each user belongs to a department. Department is characterized by name.
- Accounts – each user has an account in order to access the information system. Account is characterized by username, password, role, status, user and date of last login.

This part of the database is illustrated graphically in Figure 2. It consists of five tables with four relations between them. That is the right way for the description of these nomenclatures, because they are constant in the time chart. In this case, the storage of data in a standard relational database is more than adequate. If necessary, it is possible additional fields in any of the tables to be added, which is unlikely to have often. The illustrated database schema complies with the rules for the design of relational databases and complies with the rules of normalization of database.

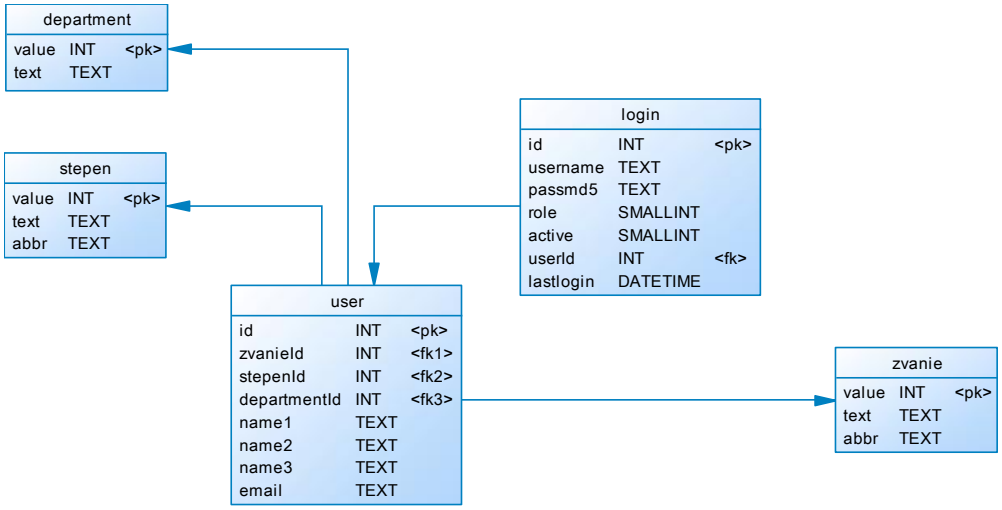


Figure 2. Database schema of the relational tables within the information system

So created tables assume the existence of standard program code for communication with the database and extraction of data from respective tables. Possible changes in the structure of these tables will surely lead to a change in the program code of the information system.

3. Dynamic part

The relational model of databases, which is used everywhere, presents each entity of the system in a separate table. Reasonable, relationships and dependencies between them are described by defining the relationship between the relational tables. Thus, changes in nomenclatures which would lead to changes in the databases of information system will cause the need for new fields in relational tables, or even new tables and relationships that will inevitably lead to changes in software solutions. Here we focus on how we can reorganize the data in the system, so changes in the database due to changes in database structure to be minimal. Let's take a look at Figure 3. We are trying to illustrate a database schema, which we will call an N-ISAV-Ref model. This schema is based on the so-called Entity-attribute-value (EAV) model [1]. EAV is a model that describes the data entities, where the number of attributes that can be used to describe them is potentially huge, but the number that actually presents attributes for a particular substance is relatively modest. Data in the EAV model is shown in three columns. Entity identifies the elements. Attribute uniquely describes a characteristic of the entity. Value contains the value of this attribute. In this sense, we will call the model at Figure 3 Nature-instance-sequence-attribute-value-reference (N-ISAV-Ref) model. In this sense, the Nature field describes the nomenclature, the Instance field is analogous to the Entity, Attribute and Value fields retain the logic of its purpose, and the fields Sequence and Reference make possible the grouping of attributes of the entity. Analyzing this model, we can say that adding attributes to existing entities will not lead to changes in the database schema, which in turn will reflect the performance of applications and any communication between individual applications. Thus, this model should be appropriate when the legislation is undergoing changes and we want to minimize the negative consequences.

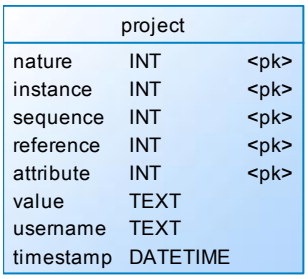


Figure 3. Database schema of the dynamic data within the information system

Table 1 provides a sample of the table of projects in the database. In such realized database, the description of the fields within the only table is implemented in the software part of the information system. Accordingly, the change in the data structure only leads to a change in the program code of the system, but not in the database schema. In this case this is a serious advantage because of the existence of dynamic data in the meaning of the system.

Table 1. Data view of N-ISAV-Ref table for a concrete scientific project

nature	instance	sequence	reference	attribute	value
1	10	0	0	0	5
1	10	0	0	5	20
1	10	0	0	10	text
1	10	0	0	20	1
1	10	0	0	30	2
1	10	1	2013	35	1
1	10	2	2014	35	1
1	10	3	2015	35	1
1	10	0	0	40	text
1	10	0	301	50	2
1	10	0	302	50	2
1	10	0	303	50	text
1	10	0	305	50	12
1	10	0	306	50	43
1	10	0	307	50	1
1	10	0	308	50	250
1	10	0	309	50	250
1	10	0	310	50	300
1	10	0	0	70	3
1	10	1	0	80	text
1	10	2	0	80	text

1	10	3	0	80	text
1	10	4	0	80	text
1	10	5	0	80	text
1	10	6	0	80	text
1	10	7	0	80	text
1	10	8	0	80	text
1	10	9	0	80	text
1	10	1	91	90	text
1	10	2	91	90	text
1	10	3	91	90	text
1	10	1	92	90	10.12.2013
1	10	2	92	90	10.12.2014
1	10	3	92	90	10.12.2015
1	10	0	0	100	12000
1	10	1	2013	100	4000
1	10	2	2014	100	4000
1	10	3	2015	100	4000

4. Conclusion

In the paper was proposed an alternative optimized schema of databases, which can be in use to the information systems with dynamic content. This schema (called N-ISAV-Ref) would be useful in terms of the easier integration between applications, which stands for quite a challenge, according to the dynamic situation in the modern World.

References

1. Frederic P. Miller, Agnes F. Vandome, John McBrewster, Entity-attribute-value Model: Data Model, Data Warehouse, Denormalization, Attribute- Value System, Linked Data, Resource Description Framework, Semantic Web, Inner- Platform Effect, Alphascript Publishing, 2010.
2. Kirilova, K., "Information Aspects of Local Administrative Policy", "Economic Alternatives", Issue 1, 2013, University of National and World Economy, ISSN 1312–7462, 2013.

Mobile Payments: Threats and Security Recommendations

Veselin Popov

D. Tsenov Academy of Economics, Bulgaria
popov@uni-svishtov.bg

Abstract. This paper presents the process of influx of mobile payments in Bulgaria, which is influenced by a number of prerequisites. A model of cooperation is presented which ensures the broad implementation of mobile payments and the specific security threats and, on its basis, the respective countermeasures are identified.

Keywords. m-payments, threats, security, security measures.

1. Introduction

The mass influx of smartphones in people's everyday life has naturally led to the emergence and rapid popularization of a new e-commerce payment method – the mobile payment. Thus e-commerce and payments can be made on the fly from a mobile phone provided that there is an Internet connection.

The purpose of this paper is to make an overview of the development of mobile payments in Bulgaria, to present a model of cooperation that ensures their broad implementation, to identify the threats to mobile payments and to outline some measures to counteract these threats.

2. Definition of Mobile Payments

According to Chae J., "Mobile payments are defined as a type of payment transaction processing in which the payer uses mobile communication techniques in conjunction with mobile devices for initiation, authorization, or completion of payment."[1]. The author also ranks the types of m-payments (see Table 1). M-payments are most often used for [2]:

- Merchandising purchases;
- Payment of accounts;
- Mobile money transfers;
- Purchase of digital products;
- Payments for transport and ticketing.

The emergence of m-payments is a natural consequence of the rapid development of e-commerce. In general, the main prerequisites leading to the widespread adoption of m-payments in our country can be summarized as follows:

- increased use of smartphones;
- increased number of people buying goods and services via the Internet;
- solutions offered for m-payments;
- increased number of the suppliers of goods and services (traders) using m-payments.

Table1. Types of m-payments

	Technology used	Purchase relationship	Charged to
Proximity	**Contactless payments**		
	• NFC • QR Codes	• C2B • B2B • P2P	• Credit card • Debit card • Prepaid account
	Hybrid payment devices (out of scope)		
	• Mobile device becomes a card reader through a hardware extension	• C2B • B2B • P2P	• Credit card • Debit card
Remote	**Message or browser payments (out of scope)**		
	• SMS • USSD • Web	• C2B • B2B	• Network bill • Debit card • Credit card • Virtual pre-paid account (PayPal)
	Application based payments (out of scope)		
	• Mobile money transfers • Virtual currencies	• P2P • C2B	• Bank account • Prepaid virtual account • Credit card

Source: Chae J., Towards a ubiquitous mobile payment solution: Exploring NFC mobile payment business models.
http://studenttheses.cbs.dk/bitstream/handle/10417/3848/johannes_sang_un_chae.pdf

In recent years there has been an upward trend in the number of people in Bulgaria who have ordered/bought goods via the Internet for personal purposes. According to NSI data [3], the dynamics in this sphere looks like this: 2004 – 1.0%, 2008 – 2.7%, 2011 – 6.7%, 2012 – 9.2%. For comparison, the average percentage for European countries is much higher; according to Eurostat [4] in 2012 it was 35% for the people who purchased goods or services over the Internet.

Payments through mobile devices in European countries are expected to reach 250 billion euros in 2014 [5].

As regards the extent of the use of smartphones Bulgaria lags slightly behind the European average. According to a survey by Google Consumer Barometer 2013 [6], the influx of smartphones in Bulgaria is 28 %, which corresponds to 2.3 million users.

The forecasts are for a significant increase in the number of users of m-payments. A survey of Jupiter Research [7] predicts that the users of mobile banking services will rise and in 2017 they will exceed the current state of 590 million USD by more than 1 billion USD. The value of the ordered goods and services via mobile devices is expected to reach 1 trillion USD. According to experts NFC-payments will be prevalent among the mobile payments in five years' time.

Solutions for mobile payments in Bulgaria

At present there are two systems of mobile payments in our country - ***mobb*** of Borica and Bankservice and ***CellumPay.***

Mobb of Borica and Bankservice is a universal platform for mobile payments. It enables smartphone users to use a variety of bank cards for their payments. Users can register multiple credit and debit cards issued by banks, which are partners of the *mobb*-initiative. Users use a special application which runs on smartphones with Android, iOS, Windows Phone and Blackberry operating systems.

CellumPay is an application for mobile payments via telephones with iOS, Android and Windows phone operating systems. It is used for paying different goods and services to the following merchants: Sofiyska voda, Euro broker, Vivacom, Green taxi, Helikon, Gladen.bg, Bulgarian Red Cross, Tip Top courier, etc.

3. Model of cooperation for the development of mobile payments

The model of cooperation for the development of mobile payments proposed by Bourreau и Verdier [8] shows the m-payment system architecture and presents the m-payment participants. We will use this model to identify the possible security threats and to outline some measures to counteract these threats.

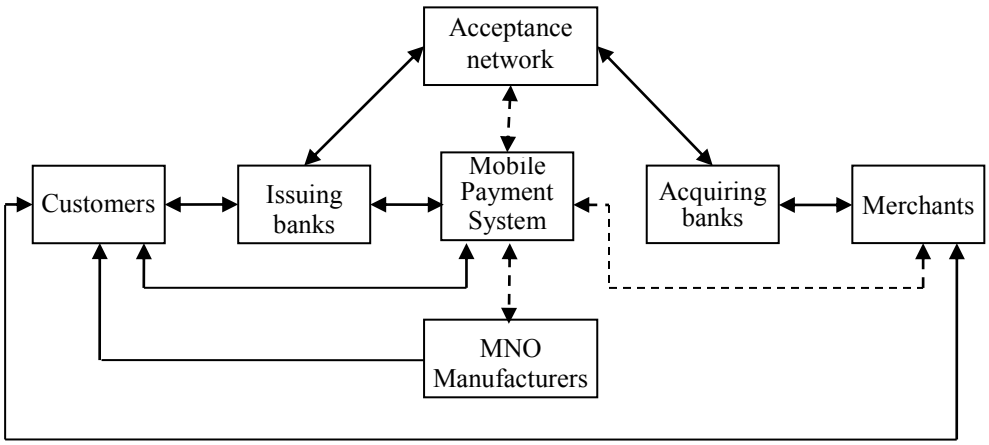


Figure 1. Model of cooperation for the development of mobile payments

Source: Bourreau, M, Verdier M. Cooperation for Innovation in Payment Systems: The Case of Mobile Payments. Communications & Strategies, 79, 3rd Q. 2010. p. 108.

In this scheme, the mobile network operator (MNO) and the phone manufacturer are those that determine and control the features, design and distribution of mobile phones, as users are typically offered phones at subsidized prices and working with their SIM cards. User accounts are controlled by the bank which supports them and mobile payment made through the networks is controlled by payment platforms such as Visa or MasterCard.

4. Threats to the type of mobile payments and measures to enhance the security of mobile payments

Mobile payments are subjected to threats that are typical of payment systems and also that are specific only for them; these threats can be summarized as follows [9]:

- Payment Tracking.
- Small Screens on Mobile Devices.
- Employees of the Payment Instrument Company
- Natural or Man-Made Disaster
- Lack of Security Implementation
- Buggy Implementations
- Dispute Resolution Inconsistency
- Proprietary Payment Protocols
- Confiscation of Money, Blocked Transactions
- Theft of Device
- Clumsy notification process after the loss or theft
- Uncertainties Regarding "Remote Wipe"
- Increasing Hacker Attention on Mobile Devices

Measures for enhancing the security of mobile payments

Ensuring protection of mobile payments is a complicated process because first of all, it depends on each one of the participants in the process (see Fig. 1). For this reason, the measures to enhance security are classified according to participants [10].

Customers must use passwords to access mobile payments applications; must not share unnecessarily confidential or personal information; must download the mobile applications only from trusted sources; must report immediately to the financial institution when a phone is lost or stolen.

Payment Brands must make sure that the security standards are processed so as to be connected to the mobile channel; must build partnerships and relationships between financial institutions and mobile operators; and must check the certification of applications and devices of the different manufacturers.

Mobile Network Operators must necessarily include software for mobile security in the devices they offer; must make sure that the smartphones for performing proximity payments are certified and meet the requirements; must check that the smartphones used in contactless payments are certified and meet the requirements; and must provide information to consumers regarding mobile security.

Financial Institutions must adapt the existing security measures to the specifics of m-payments; must adapt their back-office processes to support the new mobile channel; etc.

Vendors in the Mobile Payment Space must require that the mobile application encrypt sensitive data without relying on mobile protocols; and must control and restrict the channel for transmitting data to trusted sources, where customers can easily differentiate their authenticity.

5. Conclusion

Mobile payments are permanently entering people's everyday lives and, according to experts, this trend is likely to rise. Mobile payments security will continue to be important for their development. It is crucial for customer confidence in this kind of payments.

Providing and maintaining a high level of security in mobile payments is a task in the successful solution of which all stakeholders – manufacturers of mobile devices, mobile operators, banks, retailers and customers – are interested.

References

1. Chae J., Towards a ubiquitous mobile payment solution: Exploring NFC mobile payment business models, http://studenttheses.cbs.dk/bitstream/handle/10417/3848/johannes_sang_un_chae.pdf)
2. IE Market Research Corporation. www.iemarketresearch.com/
3. Информационно общество – Данни. Национален статистически институт. http://www.nsi.bg/bg/content/2831/%D0%BB%D0%B8%D1%86%D0%B0-%D0%BA%D0%BE%D0%B8%D1%82%D0%BE-%D1%81%D0%B0-%D0%BF%D0%BE%D1%80%D1%8A%D1%87%D0%B2%D0%B0%D0%BB%D0%B8%D0%BA%D1%83%D0%BF%D1%83%D0%B2%D0%B0%D0%BB%D0%B8-%D1%81%D1%82%D0%BE%D0%BA%D0%B8-%D0%B8-%D1%83%D1%81%D0%BB%D1%83%D0%B3%D0%B8-%D0%BF%D0%BE-%D0%B8%D0%BD%D1%82%D0%B5%D1%80%D0%BD%D0%B5%D1%82-%D0%B7%D0%B0-%D0%BB%D0%B8%D1%87%D0%BD%D0%B8-%D1%86%D0%B5%D0%BB%D0%B8
4. Статистически данни за информационното общество. European Commission, Eurostat. http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Information_society_statistics/bg
5. The 2012 Mobile Payments Market in Europe. A115. Януари 2012.
6. Google Consumer Barometer 2013. http://www.consumerbarometer2013.com/#?app=home&viewMode=0
7. Jupiter Research (http://www.investor.bg/mobilni-razplashtaniia/439/a/mobilnite-razplashtaniia--sledvashtoto-goliamo-neshto,146767/)
8. Bourreau, M, Verdier M. Cooperation for Innovation in Payment Systems: The Case of Mobile Payments. Communications & Strategies, 79, 3rd Q. 2010. p. 108
9. Adam. Have Money, Will Travel: A Brief Survey of the Mobile Payments Landscape http://blogue.priv.gc.ca/information/research-recherche/2013/mp_201306_e.pdf).
10. Gajda, M. Managing the Risks and Security Threats of Mobile Payments. http://www.pymnts.com/journal-bak/201/managing-the-risks-and-security-threats-of-mobile-payments.

Some Approaches to the Application of Imitation Modeling in Educating Students

Svetlana Vasileva [1], Alina Kulchiar [2]

[1] SHU, College – Dobrich, ZK Dobrotica 12, 9302 Dobrich, Bulgaria
svetlanaeli@abv.bg

[2] SHU, College – Dobrich, ZK Dobrotica 12, 9302 Dobrich, Bulgaria
taistaif@abv.bg

Abstract. The advantages of the simulation modeling imply its use in the training of students in various degrees. The paper offers some ideas for the application of imitation modeling learning in various disciplines such as Databases and applications, Operating systems, Computer Networks and Communications, Computer Architectures, and more. One of the classic systems of the simulation modeling GPSS World provides software tools and environment for the development of an innovative approach with practical character in the demonstration of various algorithms and visualization of work of various systems under different management protocols. Subject of this paper are the methods of inter-processors communication algorithms for mutually exclusive access to a shared resource, algorithms transactions concurrency control in database management systems, models of communication systems in different circumstances and working models of processor systems.

Keywords: Simulation modeling, process synchronization, transaction concurrency control, data transmission, electronic devices.

1. Introduction

Computer modelling is world wide spread, applicable to all spheres of the human activity from technical, technological and organizational systems to the problems concerning the human development. We can find the most effective solutions for the problems, just modelling - trying different situations. IM (imitation modeling) allows investigating queuing systems (QS) in different types of input streams and intensities of coming in of the requests of the system entrances and defining their main characteristics. It can be "played" in the time for just one trial as well as for a certain number of trials, therefore is applicable in solving optimization tasks. Moreover, the results from the modeling are decided at a random way by the character of the processes. These data are enough to receive stable statistics to be analyzed afterwards.

The simulation modeling has the following special advantages – there are only parts that are substantial for the understanding of the behavior; the model can be built before the real system with much less means; in the course of the modeling different parameters can be changed; the model renders an account of the occasional character of the processes in the real system; there is no deep knowledge in calculus necessary for the conducting of the modeling.

GPSS is a modeling language which has been used to build event discrete simulation modeling and conduct experiments on a computer [1], [2] and others. GPSS environment is used in the education process for demonstration of the work in systems already learned and

for development of models for management and work of different systems (computer systems, communication systems, Grid systems, etc.) by the students.

The main and basic application of IM is investigation of different QS but it can successfully be used also as a means of teaching especially in subjects Operation systems, Database systems, Parallel programming, Grid systems [3], etc. in the contents of which there are topics on synchronization and concurrent work of processes (transactions) to be learned.

The simulation modeling has unlimited application, this paper regards only topics from the disciplines Databases and applications, Operating systems, Computer Architectures and Computer networks and communications.

2. GPSS World – Opportunities for Modeling and Visualization of Algorithms for Synchronization

GPSS (General Purpose Simulating System) is a simulation environment with a common purpose of modeling complex discrete systems. The basic principle of work of the simulation model under the management of the simulator is the moving of the transactions from one group of blocks to another one as the transmission of the control is realized in model time. The main components of the system are transactions and blocks that present dynamic and static objects of the modeled system [1], [2].

The window of blocks and other windows in the GPSS World environment, provide excellent opportunities for observing the work of the modeled systems. By the command *Simulation Window* from the *Window* menu you can choose which windows should be open on the screen to observe the simulation:

a. *Blocks Window* - gives information about: labels and names of the blocks; number of entries in the corresponding block and the others. The window allows chronological tracking of transactions in blocks in model time;

b. *Facilities Window* - a window of single channel devices - gives information about: number / name of the device; number of inputs; rate of use, average time of residence of the transaction in the device; state of readiness; number of the last transaction occupying device; number of interrupted transaction in the device; number of interrupting device transactions; number of transaction, pending special conditions; number of transactions, pending the holding of the device;

c. *Matrix Window* – a window of the matrices - shows results in values for the total transaction time resource model;

d. Plot Window – a window of a diagram of function or of expression (not used in the presentation of the models developed because of the specifics);

e. Logic switches Window – a window of logical switches (in this case it is not used);

f. *Queues Window* – a window of queues - contains information about: number / name of the queue; maximum queue length; current queue length; total number of inputs; number of "zero" inputs; average queue length; average time of residence of the transaction in the queue and others;

g. Savevalues Window – a window of the stored values;

h. Storages Window – a window of the multi-channel devices. It contains information about: number/memory name; memory capacity; number of units free capacity; minimum quantity of used capacity; maximum quantity of used capacity; number of entries in the memory and others;

i. Table Window – a window of the tables – a diagram of the frequency distribution of the tabulated transactions.

Several windows can be open and ordered on the screen in the demonstration of a model and different aspects and elements of the modeled system can be watched in them.

3. Databases and applications. Concurrency control algorithms

Fig. 1 shows a window of simulated Centralized Two-Phase Locking with integrated mechanism of timestamps. The modeling algorithm is described in details in [4]. The combined view of the windows for demonstration and tracing of the transaction execution includes:

a. the window of facility entities bottom left of fig. 1;

b. the window of the matrices GBDA1 and GBDA2, which model the local databases for storing the two replicas of the every data element. In the windows of matrices GBDA1 and GBDA2 the students can see the synchronous update of the two replicas of a data element in the execution of the write operation over the correspondingly data element. Windows matrices GBDA1 and GBDA2 are tables with two columns first one is the number of the data element and the second one is the value of the data element;

c. the window of the stored values – in the window at the top left of fig. 1 can be observed the values of the counters: BROITR1 and BROITR2 - number of the generated transactions correspondingly length 1 and 2 data elements, BROITR - total number generated transactions, ZAVTR1 and ZAVTR2 - number of committed transactions with a length correspondingly 1 and 2 data elements, ZAVTR - total number committed transactions, RESTRT - number of restarted transactions;

d. the window of the devices called type memory – the window left in the middle fig. 1.

During the simulation of the algorithm in each of these windows can be seen the change of the values.

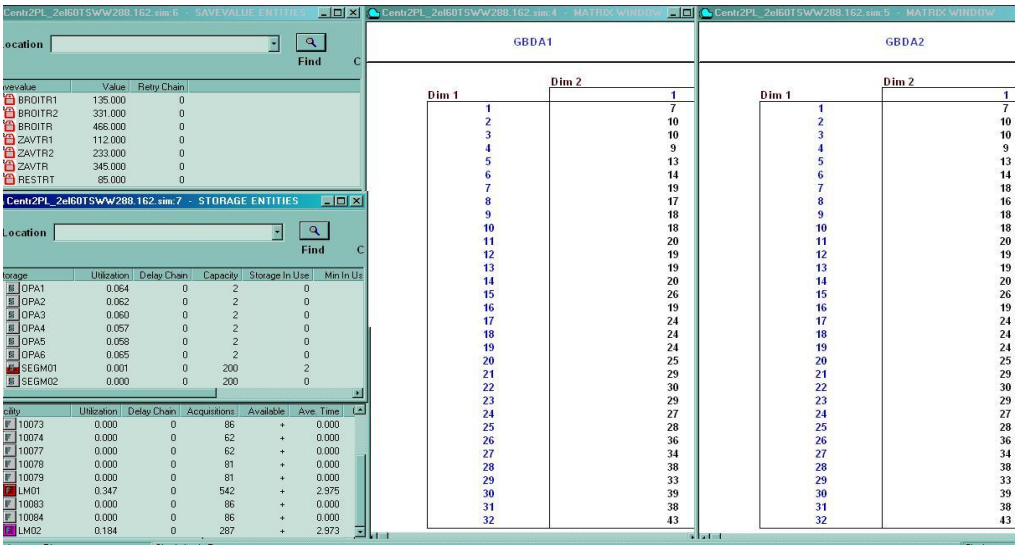


Fig. 1 Combined view of the windows for demonstration and tracing of the transaction execution

Students can observe the simulated Centralized Two-Phase Locking with integrated mechanism of timestamps and by other means arranged combined view of the windows for demonstration and tracing of the transaction execution, such as that of fig. 2, which includes (besides those mentioned above the windows of the matrices GBDA1 and GBDA2 and the window of the stored values) and the window of a table value - a diagram of frequency distribution of tabulated transactions (The frequency distribution of the committed transaction is one of the main features in the study of the database systems.).

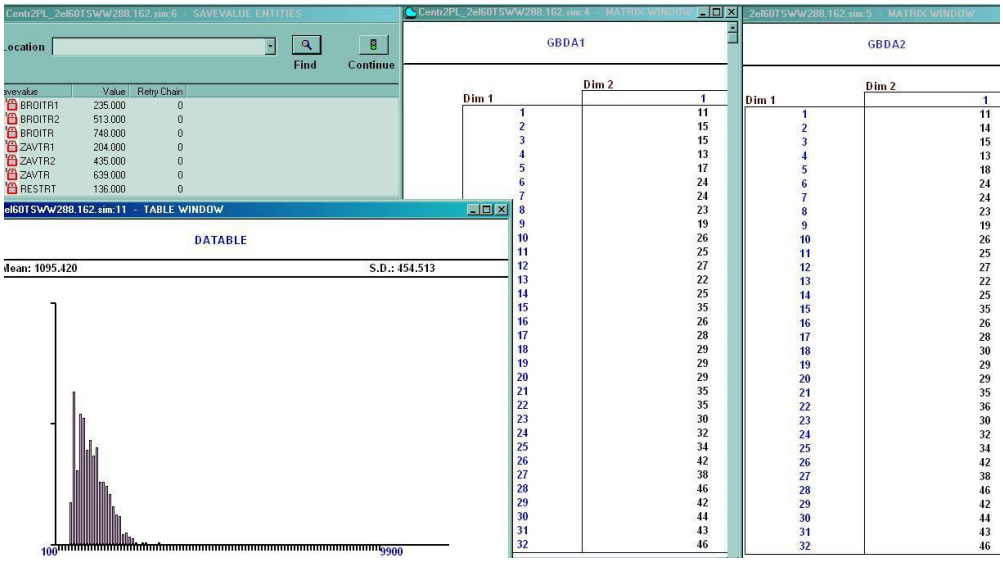


Fig. 2 Combined windows for demonstration and tracing of the transaction execution with monitoring diagram of the frequency distribution of the tabulated transactions

4. Operating systems

A fragment of the Peterson's modeling algorithm, developed in the GPSS World environment is shown in fig. 3. This model is presented in [5] and it is a GPSS "translation" of the founded in the multitude of Internet resources [6] C++ algorithm concerning the stream of processes modeled with two streams of GPSS transactions.

Fig. 4 shows such a combined window for demonstration of the execution of the Peterson's algorithm. In this window is added the window of the blocks (left of fig. 4) one can: the change of the values of the INTERES1 and INTERES2 flags (right bottom); the change in the values of elements of the TURN matrix which models the turn variable for the processes of the two streams; the window for the single channel device that models a critical section (upper right); the window of the shared resource and the actions of change in it from the two streams (under the window of the facility RESOU). In a particular execution the values written by the first stream must be different with no more than from the values that are written in the processes by the second stream.

```
NTERES1 VARIABLE 0
INTERES2 VARIABLE 0
*****************
TURN MATRIX ,2,1
OBSTR MATRIX ,2,1
INITIAL MX$TURN(1,1),0
INITIAL MX$TURN(2,1),0
INITIAL MX$OBSTR(1,1),0
INITIAL MX$OBSTR(2,1),0
*****************
GENERATE 5
ASSIGN 1,1
ADVANCE 2
TRANSFER ,ObResurs1
****************
GENERATE 5
ASSIGN 1,2
ADVANCE 2
TRANSFER ,ObResurs2
```

```
ObResurs1  SAVEVALUE
INTERES1,1
   SAVEVALUE INTERES1,1
   MSAVEVALUE TURN,1,1,0
Chaka1  TEST E
MX$TURN(2,1),0,Vliza1
   TEST E
V$INTERES2,1,Vliza1
   ADVANCE 1
   TRANSFER ,Chaka1
Vliza1 SEIZE RESOU
    ADVANCE 5
    MSAVEVALUE
OBSTR+,1,1,1
   SAVEVALUE INTERES1,0
    RELEASE RESOU
    MSAVEVALUE TURN,1,1,1
NAPUS1  TRANSFER ,NAPUS
```

```
ObResurs2  SAVEVALUE
INTERES2,1
   SAVEVALUE INTERES2,1
   MSAVEVALUE TURN,2,1,0
Chaka2  TEST E
MX$TURN(1,1),0,Vliza2
   TEST E
V$INTERES1,1,Vliza2
   ADVANCE 1
   TRANSFER ,Chaka2
Vliza2 SEIZE RESOU
    ADVANCE 5
    MSAVEVALUE
OBSTR+,2,1,1
   SAVEVALUE INTERES2,0
   RELEASE RESOU
   MSAVEVALUE TURN,2,1,1
*****************
NAPUS   TERMINATE 0
*****************
 GENERATE 5000
 TERMINATE 1
```

Fig. 3 GPSS model of Peterson's algorithm

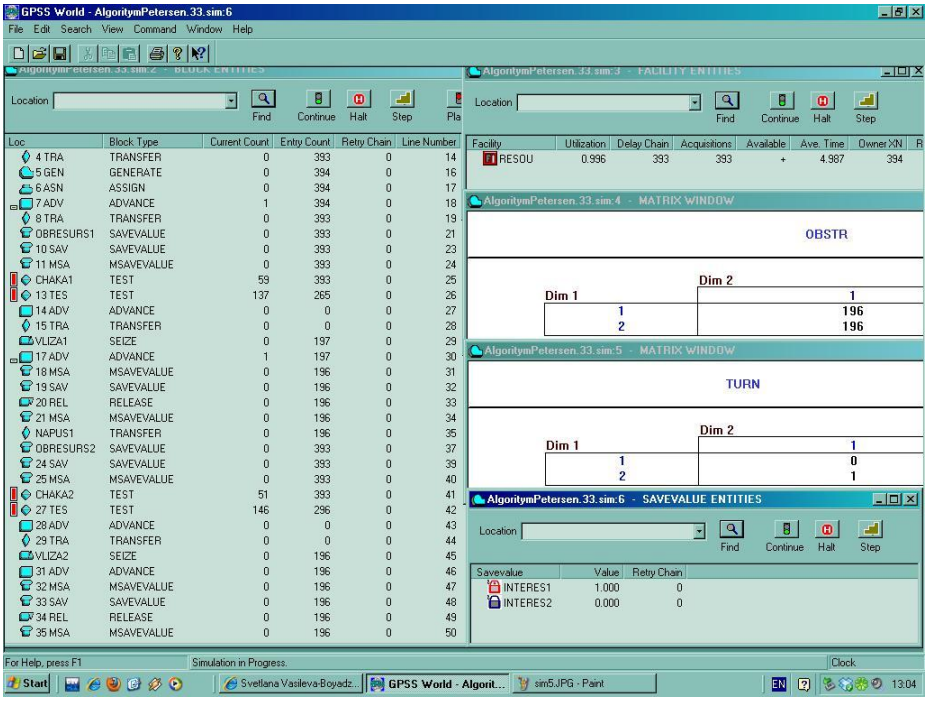


Fig. 4 A combination of windows in the demonstration of Peterson's algorithm

5. Computer architectures

Fig. 5 shows the structure scheme of simple electronic devices. The structural scheme and its presentation as a queuing system (fig. 6) are presented in [7]. The device consists of an input buffer, main channel service and backup channel.

The GPSS World code of the model described in [7] is the following:

```
GENERATE 8, 4   ; generation transactions
QUEUE QRAM   ; entrance drive
QUEUE QN        ; input queue to the main channel
SEIZE OSKANAL; occupy the main channel
DEPART QN        ; to turn off the main channel
ADVANCE 7, 3      ; processing delay
RELEASE OSKANAL; release the main channel
DEPART QRAM        ; drive off
OUT TERMINATE      ; transactions output from the model
GENERATE 170, 30, 1; generation of "bad" transactions
PREEMPT OSKANAL, PR, OUT; Capture both the main channel
SPLIT 1, REZ             ; Send a copy of a backup channel transaction
ADVANCE 20, 7          ; delay the restoration of the main channel
RETURN OSKANAL   ; release the main channel
TRANSFER, OUT       ; send transaction to block TERMINATE
REZ QUEUE 1      ; take a turn
ADVANCE 1, 1      ; delay to start the backup channel
DEPART 1           ; leave the queue
QUEUE 2            ; input queue to the backup channel
SEIZE RKANAL   ; occupy backup channel
DEPART 2           ; turn off to the backup channel
ADVANCE 7, 3      ; processing delay
RELEASE RKANAL   ; free backup channel
TRANSFER, OUT       ; send transaction to block TERMINATE
GENERATE 7200       ; transactions generation through 7200 model units
TERMINATE 1; Transaction output from the model and reducing the counter value by 1
START 1          ; setting the initial value of the counter in one unit
```

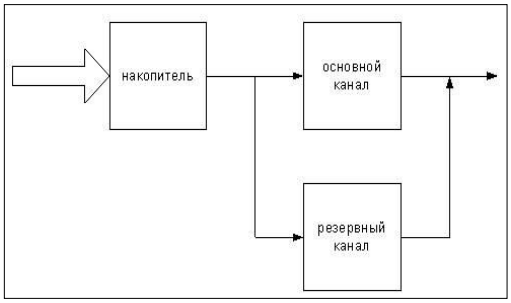


Fig. 5 Structure scheme of electronic device by [7]

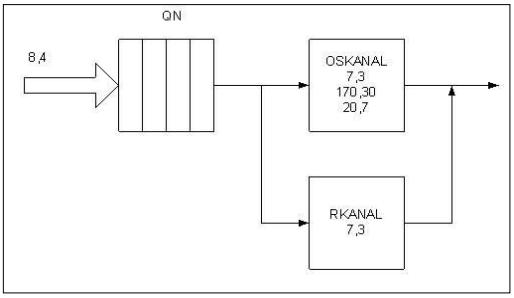


Fig. 6 Description of the electronic device as a queuing system

Fig. 7 shows such a combined window for demonstration of the work of the electronic device. Right of fig. 7 the window of the blocks where the students can observe the transaction moving in the model. Top right on the fig. 7 shows the windows of the facilities, where you can see how many transactions are serviced by the main channel, and the backup channel. Bottom right of the fig. 7 shows the window of the queues, which are organized on the model: in front of the electronic device, in front of the service channels, in front of the main channel and in front of the backup channel.

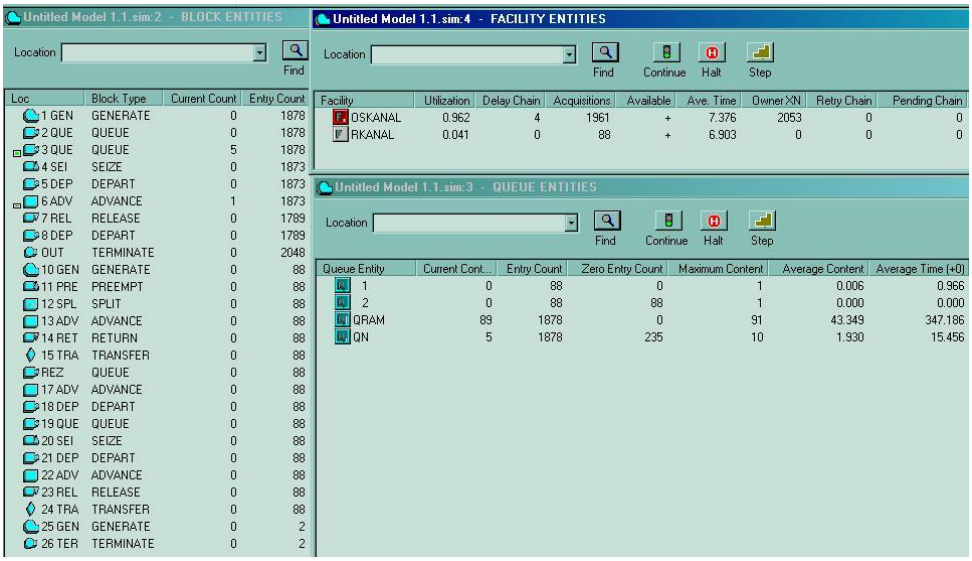


Fig. 7 A combination of windows in the demonstration of electronic device work

Fig. 8 shows the report with the simulation results. If the windows of fig. 1, fig. 2, fig. 4 and fig. 7 show the work of relevant systems "live", "on the move" the reports with the simulation results show the work of the system "in summarized form." Windows accounts from the simulation always appear after the end of the simulation and can be successfully used to analyze the performance of the studied system without worry about the time. The reports of simulation can be saved as separate files and open later (and elsewhere) and examined in detail. Fig. 8 shows the final look of the windows. Fig. 8 shows the final look of the windows included in the combined window of fig. 7.

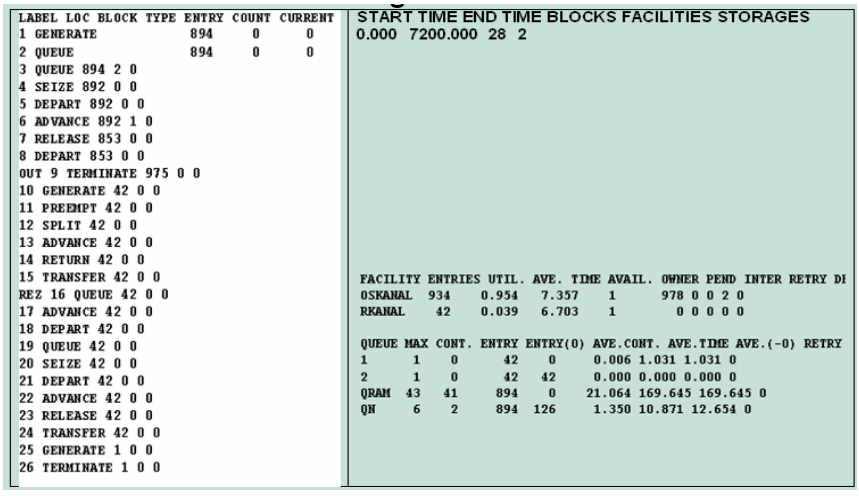


Fig. 8 Electronic device simulation results with listing

6. Computer networks and communications

Fig. 9 shows the structure scheme of a system for the transmission of data. The structural scheme and its presentation as a queuing system (fig. 10) are presented in [7]. The system from the fig. 9 has 3 points: A, B and C. It has 5 lines: AB1, AB2, AB3, BC1 and BC2. These components present in the model (fig. 10): the points by the queues; the lines by facilities; and the data packets by transactions.

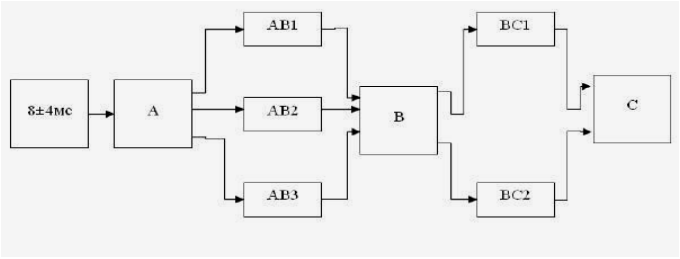


Fig. 9 The structure scheme of a system for the transmission of data

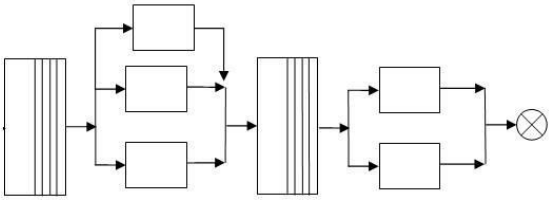


Fig. 10 Description of the system as a queuing system

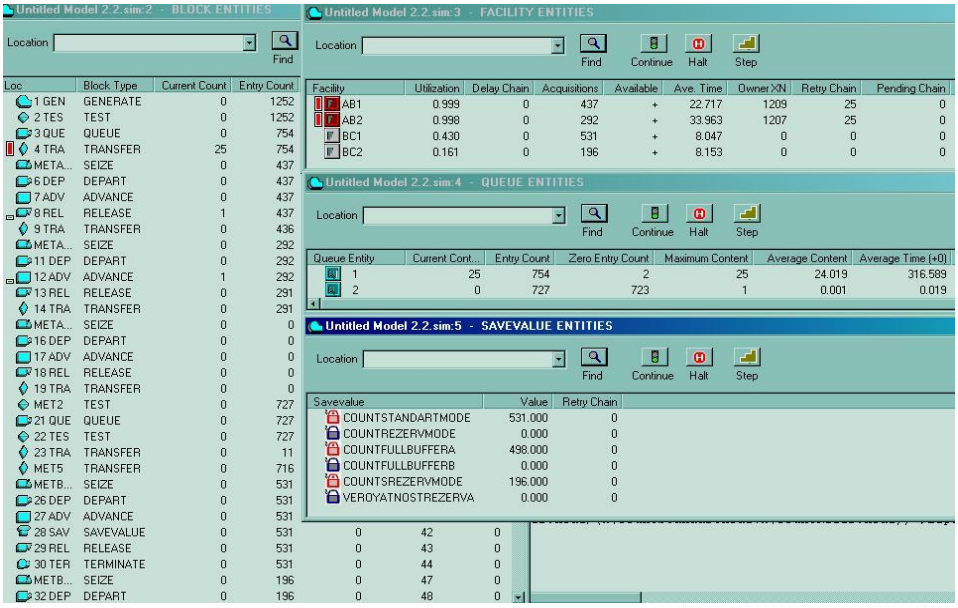
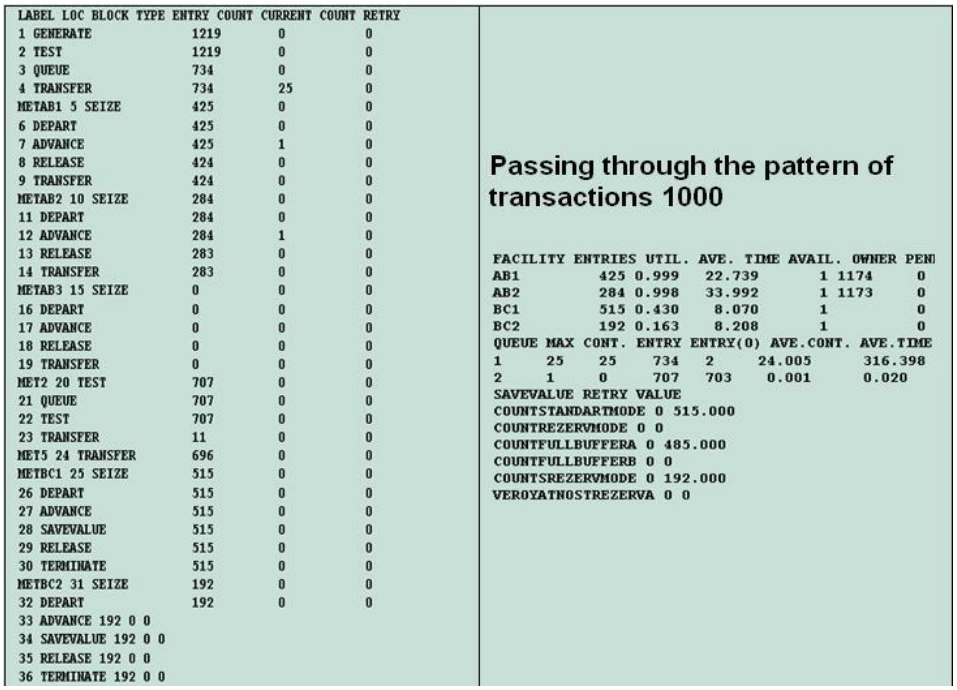


Fig. 11 A combination of windows in the demonstration of system for the data transmission work

As was mentioned in the previous section, except the block windows, the facility's windows, the save values windows, etc., when running on the system can be seen "live" for educational purposes can be used with simulation reports the final results of simulations. Fig. 12 shows fragments of the report from executing model described in fig. 9 - fig. 11.



LABEL LOC	BLOCK TYPE	ENTRY COUNT	CURRENT COUNT	RETRY
1	GENERATE	1219	0	0
2	TEST	1219	0	0
3	QUEUE	734	0	0
4	TRANSFER	734	25	0
METAB1	5 SEIZE	425	0	0
6	DEPART	425	0	0
7	ADVANCE	425	1	0
8	RELEASE	424	0	0
9	TRANSFER	424	0	0
METAB2	10 SEIZE	284	0	0
11	DEPART	284	0	0
12	ADVANCE	284	1	0
13	RELEASE	283	0	0
14	TRANSFER	283	0	0
METAB3	15 SEIZE	0	0	0
16	DEPART	0	0	0
17	ADVANCE	0	0	0
18	RELEASE	0	0	0
19	TRANSFER	0	0	0
MET2	20 TEST	707	0	0
21	QUEUE	707	0	0
22	TEST	707	0	0
23	TRANSFER	11	0	0
MET5	24 TRANSFER	696	0	0
METBC1	25 SEIZE	515	0	0
26	DEPART	515	0	0
27	ADVANCE	515	0	0
28	SAVEVALUE	515	0	0
29	RELEASE	515	0	0
30	TERMINATE	515	0	0
METBC2	31 SEIZE	192	0	0
32	DEPART	192	0	0
33	ADVANCE 192 0 0			
34	SAVEVALUE 192 0 0			
35	RELEASE 192 0 0			
36	TERMINATE 192 0 0			

Passing through the pattern of transactions 1000

FACILITY	ENTRIES	UTIL.	AVE. TIME	AVAIL.	OWNER	PEN
AB1	425	0.999	22.739	1	1174	0
AB2	284	0.998	33.992	1	1173	0
BC1	515	0.430	8.070	1	0	0
BC2	192	0.163	8.208	1	0	0

QUEUE	MAX	CONT.	ENTRY	ENTRY(0)	AVE.CONT.	AVE.TIME
1	25	25	734	2	24.005	316.398
2	1	0	707	703	0.001	0.020

SAVEVALUE	RETRY	VALUE
COUNTSTANDARTMODE	0	515.000
COUNTREZERVMODE	0	0
COUNTFULLBUFFERA	0	485.000
COUNTFULLBUFFERB	0	0
COUNTSREZERVMODE	0	192.000
VEROYATNOSTREZERVA	0	0

Fig. 12 Simulation results with listing demonstrating the work of system for the data transmission

7. Conclusions

The presented models developed in the environment for simulation modeling GPSS World, are informational only – serve for visualization of concurrency control (in Database subject), inter-processor communication (in Operation systems subject) and not only. Opportunies for including simulation models in the education process for demonstration of the algorithms and not only, are countless. Further the developed models and the simulation modeling can be improved and therefore be more complex and effective. GPSS World with all its potential can promote the quality of students' education, not only as it is shown in this paper, in others subjects as well (Computer Network and Communications and Parallel Programming).

References

1. Minuteman Software. General Purpose Simulation System World, www.minutemansoftware.com 2012.
2. Modeling of systems and networks based on GPSS. 2008. http://en.radioland.net.ua/contentid-160-page1.html.
3. http://ci.fmi.uni-sofia.bg/VasilGeorgiev/datagrid/grids13.pdf.
4. Vasileva, S. 2013. An algorithm modelling centralized two-phase locking with integrated mechanism of timestamps. Journal of the Technical University Sofia, branch Plovdiv. "Fundamental Sciences and Applications", Vol. 19(1), International Conference Engineering, Technologies and Systems TechSys 2013, 237-242.
5. Vasileva, S., K. Kapinchev. Some Simulation Models of Process' Synchronization for Educational Purpose. // International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE - 2012), October 5 -6th, 2012, UNWE, Sofia, 2012, Bulgaria, pp.474-483.
6. http://www.bg-informatika.com/2011/07/22.html.
7. Computer modeling. http://www.intuit.ru/studies/courses/643/499/lecture/6490.

The Classical Education from the Perspective of ICT

Valentina Terzieva, Petia Kademova-Katzarova, Rumen Andreev

IICT - BAS, Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria
{valia, petia, rumen}@isdip.bas.bg

Absract. Modern technologies provide tools for development infrastructure platforms for implementation of e-Learning practices and digital libraries. The paper presents alternative forms of access to education and ICT assisted new methods for adoption, perception and acquiring knowledge. An application of ICT in classical pedagogical tools to achieve an educational process of higher quality is discussed. The new tools offered by ICT enable a completely new design of learning resources, enriched with a variety of colours, shapes, animation, simulation, etc. and offer a possibility for expanding the volume of presented learning material and information. Thus, even the most boring subject matter can be made interesting and understandable and the performance can be adapted to the specific characteristics of each learner. In this way all students regardless of their learning habits and style, as well as personal preferences can choose the most effective way for them to master the subject matter. ICT supported pedagogical approaches that increase the effectiveness of teaching and training, and encourage students are presented. Those are very useful for students with special education needs too; so that the ICT based teaching tools can be widely used and are suitable for all learners. The implementation of the classical teaching methods through techniques enriched with the ICT-offered novel agents is a major challenge for teachers in all phases of education - primary and secondary school, as well as at the university.

Keywords. ICT based teaching, technology enhanced education, e-Learning resources, adaptation, learning style.

1. Technology in Education

The rapid development of modern ICT and their application in – one can say all areas of life, from everyday wear to the latest achievements of science – imperatively require their use in education. This process has two sides:

1. the use of ICT as a learning tool and
2. the studying their capabilities and their effective use.

In the first case the trivial implementation of ICT (interactive whiteboard – a substitute for a black one; equipment for each student place with computer / laptop / tablet – a substitute for conventional textbooks and notebooks) is not enough. Publishers now produce variants of textbooks for the Internet that are interactive to some degree. Universities provide students with lectures and literature in the local network. Often the assigned tasks and projects for each student are set there too. Despite these undeniable good innovations in education, quality and efficiency of teaching, exercising and testing of acquired knowledge can be significantly improved. ICTs provide information on almost all issues in a variety of volumes and forms: text, images, movies, animated movies, animated illustrations, audio, life-sessions, augmented reality and even games, and thereby the content may be at a different level – from strictly scientific, presenting the latest research on the

problem to a simple popular explanation, suitable for children or amateurs. Besides published in the Web teaching materials, amateur ones can be found too and very often the last are more readable and easier to acquire, because not constrained by the rules of "formal education". In the ocean of information available on the Internet, the teacher can find and construct the learning material in a very interested and engaged manner, taking into account the special features of his / her audience. The information flood sometimes makes it difficult for even the presentation of the material. And that is already one of the issues in question in the second mentioned case – the potential of ICT and their effective use [1].

As new devices and applications using ICT are appearing continually, learning how to handle them is extremely important. Even the use of "ordinary" products, such as Word, PowerPoint or smart phone is not always easy because there are subtleties known by relatively few people, while the products are incessantly updated. Similarly is the problem when extracting information from the Web. Almost each of the participants adds new information, so it is difficult to pre-screen the right content, and to check its authenticity. It requires certain knowledge and skills to actually find what interests us, and use the minimum of time to remove the unnecessary.

Implementation of ICT in education can improve learning because of the rich capabilities possessed:

• access to a vast amount of information;

• use of various forms to present information;

• combination of different methods and educational tools to increase the efficiency of the educational process (visualization, explanations from several different perspectives, "tailored" learning paths);

• creating skills to operate with new tools and ability to follow their developments, adopt and implement them.

2. ICT- Alternative Forms of Access to Education

Besides the variety in the presentation of learning material in educational institution, ICT provides opportunities for alternative form of education. Many universities (especially in the USA) have distance learning programs, which differs greatly from the hitherto used form (supply of literature – a few part-time courses – exam). Now the whole study material is available in electronic format (books, lectures, assignments, tasks, projects). ICT allows for consultation (in synchronous or asynchronous mode), drafting, carrying out classes, seminars, problem solving and even giving exams.

There are other forms of education, e.g. autonomous learning. Usually the learner deals with a particular problem on amateur basis, but gradually enters the selected area and looking for sources of information on the problem, qualifies to the extent that he/she can withstand the tests required to obtain an official certificate. This applies not only to scientific fields, but also for certain skills. Here are some of the most common types of education relevant to the present:

- university programs – whole degree program (NBU) or individual courses (free at many universities in the USA);
- company training – Microsoft & Cisco Academy, Telerik Academy;
- interactive lessons & workbooks;
- thematic forums for guidance and assistance;
- platforms with content, which is not necessarily structured and of quality;
- additional qualification, which expands the area of institutionalized education.

But, as already mentioned, to be successful training is therefore necessary to create skills for working with web information. Finding, extracting, selection of reliable information are important in training in the last three types. In the first three cases, the training materials are provided from certified sources, which subsequently, after appropriate testing, may issue an official document.

3. Pedagogical Methods

By "Pedagogical methods" are indicated the principles and methods used for teaching. The choice of method depends mostly on the type of information or skills that are taught, and are influenced by the knowledge and abilities of the students at that moment. Basic teaching methods can be divided into three types and several sub-types (Figure 1). Usually teaching is carried out not only according to a single method, but a mixture of two or more types is implemented:

– **Informative.** This method is performed by oral or written presentation of facts, definitions, explanations, evidence. Reading or listening to the material (from a book, newspaper, magazine, lecture, radio, TV) is an example of the transmission of knowledge in informative way. Training is done through "instructions" and explanations. This is the "classical" didactic teaching, thereby giving oral or written explanations rather than demonstrations of matter. Often this method is referred to as lecturing. The key elements of informative teaching method are messages and symbols. Written and verbal presentation of information can be enriched with illustrations (diagrams, schemes, photos, drawings) for illustrative purposes.

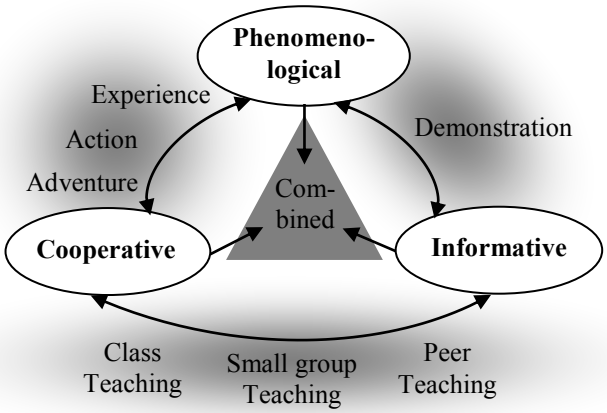


Figure 1: Basic pedagogical methods

– **Phenomenological.** This method is based on the demonstration of phenomena, events, processes. Performing experiments at school (physics, chemistry, biology, etc.), observation of natural phenomena, of events in everyday life, implementation of self-designed experiments lead to gaining knowledge in this way, i.e. knowledge is built on events and is absorbed and understand through senses and emotions. There are four basic types of phenomenological approach:

- ***Demonstration*** is a process of teaching through examples. It is a combination of visual evidence and related considerations. E.g. memorizing a list of "dry" facts is detached and impersonal activity, while the same information presented through demonstration is personally experienced. Demonstrations help to increase student interest and enhance retention because they provide links between facts and real-world applications.
- ***Immersion in a problem*** focuses on learning for the individual. Teaching is done by monitoring the target object, thus the learner makes discoveries with knowledge firsthand, instead of listening to or reading about others' experience.
- ***Action*** – (learning by doing) is one of the most powerful tools for learning. It requires self-initiative, strong intention of learning and active learning phase [2]. Learners acquire knowledge through real actions and practice, rather than the traditional way. The examination of a problem "important issues" gives rise to reflections on the desired goal, the development of strategies, developing action plans and implementation for performing the steps comprising the application plan [3].
- ***Adventure*** is a hybrid approach to distance learning [4]. It allows learners to explore real-world problems through authentic learning experiences within collaborative online learning environment. This approach comes from learning by doing [5] and includes educational activities with authentic experiences of researchers "on the spot". For example, the curriculum, the experience and observations of scientists, the online opportunities for interaction between participants offered synchronously so students are able to make the connection between what happens in the real world and their research, and then reflect on the events and provide possible solutions to the raised issues [6].

- Cooperative. In cooperative learning method knowledge and skills are formed in a group. A difference has to be made between education in everyday life and in an educational institution. Very often the first aspect is not mentioned at all, because the emphasis is on educational programs. But actually this is the real initial training, starting from the first moment of life and continuing until its end. In this case knowledge and skills are formed in the socio- cultural environment in which the individual lives (family, ethnic group, thematic forums, discussion groups, etc.) and "is transmitted" between the group members. This are shared knowledge, skills and experience. Learning the mother tongue, social skills, information about hobbies are examples of obtaining knowledge in this way. Cooperative learning methods are usually equally effective for all levels of ability and all ethnic groups; they increase self-confidence and self-esteem [7].

- ***Teaching in class*** is a very common method in educational institutions (school, university). But the formal learning process requires taking positions (roles) – the teacher and the student – and thus establishes a distance between them while in everyday life the contact is direct.
- ***Teaching in small group*** – the class is a large group and it is more efficient to form smaller groups (2 – 6 participants) to perform certain tasks. Unlike individual learning, which can be competitive, students in conjunction benefit from common resources and skills (searching information among themselves, appreciate each other's ideas, monitor work, etc.) [7]. The role of the teacher changes: from "only" a source of information it becomes an assistant in training [8, 9] or moderator. Everyone is successful, when the group succeeds. Collaboration allows students to actively participate in the learning process and engage in dialogue with each other and listen to other points of view. It establishes a personal relationship between students and the

subject of study and helps them to think less biased. The small group can be formed real or virtually.

- ***Peer teaching*** is a method of training, where students assume the role of teacher and teach their peers. First they need to study and understand the subject well enough. In the teaching process they gain self-confidence, improve their language and communication skills.

The above mentioned learning sub-methods by action and by adventure (phenomenological type) can also be assigned to the cooperative method of teaching. Rarely can be met a "pure" method of teaching, mostly it is a combination of several ones (Table 1).

Pedagogical methods			Sub-methods	ICT Tools	
Informative – Minding: read, hear watch			Didactic teaching	**text**, **audio** (music, recitation), **image** (chart, table, picture, photo, scheme, cartoon)	GAMES
		Phenomenological – Making: observe, reflect, empathize	Demonstration	**image** (animation, photo, video clip, comics, movie), **audio** (sound effect, music, recitation)	
			Immersion in a problem	artistic **reading**, artistic **picture**, **film**, **theatre**, **storytelling**	
	Cooperative – Socializing: exchange information		Action	making **virtual experiment** in **virtual labs** through **simulation** of real processes, (the appropriate reactions are simulated too), drawing, painting, singing, role playing, making music or audio effect etc.	
			Adventure	remotely participation on real processes – **expeditions**, **experiments** – via life Internet connection	
Informative – Minding: read, hear, watch			Small group teaching	all mentioned tools for didactic teaching, and phenomenological teaching; **exchange of information** between students – e-mail, social networking, chat, network game, carrying out joint project	
			Peer teaching	all mentioned tools for didactic teaching, demonstration, immersion in a problem and action; **exchange of information** between students – e-mail, social networking, chat	
			Class teaching		

Table 1: ICT implemented in basic pedagogical methods

4. Application of ICT Tools in traditional education

There are two equally important aspects of any education process – knowledge and effective teaching methods. Both of them are significant for the education of high quality. It is obvious that there isn't any strong formula what teaching methods are useful for any subject. The traditional teaching methods are necessary to be changed so that to bring proper effectiveness in the student's knowledge acquisition, understanding, and skill application in order to give proper studying experience to learners. The most important criteria for changes

should be the progress of the students. The variety should be introduced in the subject teaching also for motivating purpose.

ICT has many opportunities for technical support of all pedagogical methods of teaching [10]. A competent teacher should use appropriate educational technology tools in order to ensure a better teaching-learning process. Besides the trivial presentation of information in the form of text (of graphic or audio type), static images (schemes, graphics, paintings, drawings, photographs) and dynamic representations (movies, animations), visual and audio records of significant art works and performances, the new technical tools can make us witnesses of recent and historical events, evolution of scientific developments, natural phenomena and processes. All the knowledge and documents, all events, opinions of individuals and groups can be shared and accessed. In this regard ICT and especially multimedia is very useful. Multimedia text learning resources allow interactivity, automatic searching and use links to navigate to external information. In fact, multimedia learning resources enable multisensory teaching method that engages two main senses – hearing and seeing, in addition to doing. Well-known learning pyramid confirm that students retain about 70% of what they see, hear, and do simultaneously, i.e. when use complete multisensory learning resources, so they are highly effective [11].

The simulation of different processes can be used for scientific purposes as well as for learning – demonstration, experimentation and skills training. Particularly interesting is the use of the "adventure method" in the educational process – e.g. remote real time participation of students in scientific research "on the field" (program "Jason" – diving in the ocean – and classroom "Connect" generate a comprehensive curriculum and learning objectives related to practical experience by engaging students in their approach "learner chosen route"; "GoNorth! Adventure Learning Series of circumpolar Arctic dog sledding expeditions" explores themes such as sustainable development, environment, science and traditional cultures).

ICT enable greater and better mixing of the various teaching methods in order to achieve a high quality of training. Modern computer games have a special place in this process [12]. The game has always been the best way to study; in it is always an entertaining element that largely can provide motivation. Games enable practice knowledge that feels like play. An interesting feature is also the fact that the game perfectly integrates all possible methods of learning.

Many researches have proven that usage of various methods of teachings based on technology actually affect knowledge acquisition and bring about students progress [13]. Usually learners enjoy learning through multisensory resources, especially in case of serious games. That is why ICT is such a powerful tool that changes to great extent contemporary teaching-learning methods.

The students need a much wider access to web-based systems for qualification and retraining, which can be done in parallel with the traditional "classroom" education. Today's computer generation needs training in another dimension, and this will inevitably lead to the rise of Web-based systems for education and training. In the near future students usually will learn several areas of studies simultaneously, which will become increasingly possible again for distance learning and the introduction of flexible forms of teaching and assessment.

ICT will facilitate lifelong learning, because what has been learned in the first years of life will not be sufficed for the next years due to continuous technological revolution. Hence the enormous challenge to educational institutions have to create a system for training that will enable everyone to be motivated and have access to the system for (re)training. A new technology provides many more opportunities and tools for training and qualification.

Knowledge and skills acquired informally becomes increasingly important for employers. Therefore there should be a tool for the recognition of those knowledge and skills.

5. ICT for Students' Motivation

Modern technology can help everyone to learn almost whatever they want, but they are not enough for learning how the learned matter can be applied into practice. The most significant advantage of e-learning is that it allows much more easily access to collections of huge amount of knowledge and learning resources and in addition presenting them in such way that stimulates learners to be eager to search more and more information.

What is the secret of successful training – according to psychologists the key is motivation. Students have to take pleasure in what they do, in order to develop their creativity and talents. The leading motivation for acquiring knowledge is the understanding that this knowledge will be useful for you. At the outset of education new students have to see examples of their predecessors who reached the peaks of science and intuitively begin to seek to emulate them. This is the practical manifestation of the good example. ICT can bring much about this – with information on school/ university sites about most impressive students' achievements, with short clips showing successful careers of former students or even by providing opportunities for chat and/or audio/video conference with famous and/or prosperous alumni.

Learning platform, applied methods of teaching & evaluation and ICT tools for cooperation between teachers and students are crucial to the formation of positive attitudes and motivation to learn in e-learning environment. Teachers who provide less cramming and memorizing lessons and encourage students to develop their thinking and learning to use acquired knowledge in various situations provoke continued interest in the learning mater. Understanding the connection between theory and the reality is much more interesting and motivating for most students than learning abstract ideas by heart or concepts without stating clearly their practical application. At that point a multisensory teaching approach enriched with ICT tools for illustration and visualization of real processes and events is useful to the great extend. Students, who utilize educational material focusing on its practical aspects, have a deeper interest in science and confidence to tackle real projects and problems. Contemporary ICT-based communication tools incorporated in e-learning platforms (chats, e-mails, forums, etc.) can help much in establishing more informal relationships between teachers and students that will benefit both sides. Variety of applied ICT-based tools (audio/ video conference, remote lecture presentation, films, simulation of real processes, augmented reality, etc.) in traditional teaching methods make students more intrigued and interested in presented learning matter.

Another aspect of e-learning that contributes to the motivation of students is enabling relatively autonomous educational process. It is well known that people perform best when they are given autonomy, opportunity for mastery, and the belief that their task is meaningful. Learning with self paced speed, in a convenient time, on a self selected courses and/ or using preferable teaching method allows intrinsically-based motivations to increase.

Teachers should understand the unique personality types of their students to keep them interested in learning, inspired and productive. The results can be amazing if students are offered the right incentives – the ones that influence their personality and develop their brain – to make them mentally and emotionally devoted in doing their best. Money is not the universal motivator and so the excellent and motivated students have to be "rewarded" personally, for example to have choice to compose their own learning plan or to select project topic, free courses, etc. The well known work principle in Google "20% time, in

which employees may spend one day a week on whatever they want" is a brilliant example of how intrinsically-based motivations thrive and result in useful things such as Gmail, Google News, AdSense (a program that allows serving automatic text, image, video, or advertisements that are targeted to site content and/or audience) and Orkut (social networking website). Similarly, students from upper grades could be allowed to choose for example 20% or more of their non obligatory e-learning courses even from another department or university, what is very easy in web-based teaching.

ICT-based teaching methods allow students to learn on their own. The key is to make students self-motivated. There is never just one right way to do something and e-learning resources enable students to develop their own ways. Students motivate themselves in their own way, they do not have to be told how to do things, but only what to do and let them reach their own results. ICT-based platforms should provide tutors tools for guiding by quick and adequate feedback.

6. Conclusion

Having in mind the above mentioned, it is obvious that the integration of ICT in education may contribute to reducing the difficulties and obstacles of the students in acquiring knowledge concepts. In addition, the appropriate ICT-based learning courses appear to be a modern approach to self-teaching. People should educate themselves throughout their lives and promote the best of its uniqueness and experience. This can be achieved by appropriate ICT-based methods and forms that intensify educational impact and interaction; stimulate active and interactive learning and contribute to continuously improvement of education quality.

The advantages of integration of modern ICT in the education process are:

- Interactivity;
- Easy personalization of learning courses through ICT tools;
- Dynamic and adaptive learning resources;
- Multimedia presentation of content;
- Instant access to the vast amount of information available on the World Wide Web;
- Live connection to events and processes and
- The possibility of a combined use of all these properties.

References

1. Vorhaus J. (Eds.) Effective Teaching and Learning Using ICT, NRDC, London 2007.
2. Moon, J.: A Handbook of Reflective and Experiential Learning: Theory and Practice. London: Routledge Falmer. p. 126, 2004.
3. Marquardt, M., Leonard, H. S., Freedman, A., & Hill, C.: Action learning for developing leaders and organizations: Principles, strategies, and cases. Washington, DC: APA 2009.
4. Moos, D., and Honkomp, B.: Adventure learning: Motivating students in a Minnesota middle school, Journal of Research on Technology in Education, 43 (3), 231–252, 2011.
5. Slavin, R. E.: Cooperative Learning. New Jersey, Prentice-Hall, 1990.
6. Chiu, M. M.: Adapting teacher interventions to student needs during cooperative learning. American Educational Research Journal, 41, 365–399, 2004.
7. Siltala, R.: Innovativity and cooperative learning in business life and teaching. University of Turku, 2010.
8. Brown, H., & Ciuffetelli, D. C. (Eds.): Foundational methods: Understanding teaching and learning, p. 507. Toronto: Pearson Education, 2009.
9. C455 Class Notes, homepages.ius.edu
10. Brown J. S., How to Connect Technology and Content in the Service of Learning, The Chronicle of Higher Education, Vol. 55, issue 8, pp. A120, 2008.
11. Wood E. J., Problem-Based Learning: Exploiting Knowledge of How People Learn to Promote Effective Learning, Bioscience Education, BEE-j Vol. 3, May 2004.
12. Terzieva V., P. Kademova-Katzarova: Computer games in multisensory education, Proceedings of International Conference Automatics and Informatics'12, October 3 -5 2012, pp. 365 – 368.
13. Ruth H. Maki, William S. Maki, Michele Patterson, and P. David Whittaker, Evaluation of a Web-based introductory psychology course: I. Learning and satisfaction in on-line versus lecture courses, Behavior Research Methods. Instruments, & Computers 2000, 32 (2). 230-239.

Cyber Security Risk – An Important Business Continuity Planning Issue for Business Organisations

Asen Bozhikov

D. A. Tsenov Academy of Economics, 5250 Svishtov, Bulgaria
asen@uni-svishtov.bg

Abstract. A business continuity plan is essential for every business organisation. Its role is to guarantee the continuity of critical business operations, minimum business downtime and the fast and smooth recovery of key business activities in case of an occurrence of any disruptive event, no matter what the source of the event is – hurricane, fire, espionage, data theft, power outage, hard disk failure, etc. To be successful, a business continuity plan should address all the potential risks connected to mission critical business operations which should be identified, assessed and effectively planned for. As information technologies become more ingrained in business organisations' operations and communications with customers and suppliers, the rise and spread of cyber security threats will continue to multiply, leading to devastating results for business organisations. In order to deal with this, companies need to fully understand how cyber threats can affect their mission critical business activities. Thus, cyber security risks should be considered as an important issue when it comes to business continuity planning.

Keywords. Business continuity planning, cyber security risk, cyber security.

1. Introduction

Business continuity concerns the maintenance of the availability of all key business resources which are necessary to support essential business operations. Therefore, we can say that business continuity planning is connected with the development, implementation and regular update of frameworks, programmes and policies which have a main goal of preventing possible business disruptions caused by expected or unexpected events. An essential part of business continuity planning is to conduct risk assessment and business impact analysis to identify the possible threats which could cause business interruption.

There are many threats that could cause potential damage to critical business operations, not only to the operations but also to people, data, the IT infrastructure etc. Thus, business continuity planning is connected with the risk management strategy of the business organisation and it should cover a subset of operational risks.

It requires taking proactive action against the two main groups of disasters which could harm the organisation: natural and man-made disasters. Among the threats in the second group there is one that is becoming more and more important to consider – cyber security risk. The reason for this is the total penetration of information and communication technologies (ITC) in business. Nowadays data and information are vital for every business organisation. Together with the intensive use of ITC they serve businesses of all sizes in that they give competitive advantages on the global market. Protecting data and information systems and keeping them secure is a major task for the organisation.

The purpose of this report is to review the significance of cyber security risks and the effects on business organisations as this surely transforms into an essential business continuity planning issue. As the director of FBI Robert Mueller said, "I am convinced that there are only two types of companies: those that have been hacked and those that will be. And even they are converging into one category: companies that have been hacked and will be hacked again" (This quotation is part of the speech of the FBI Director on RSA Cyber Security Conference which took place in San Francisco, California in 2012. The full speech could be found here: http://www.fbi.gov/news/speeches/combating-threats-in-the-cyber-world-outsmarting-terrorists-hackers-and-spies).

2. Concerns over Cyber Risks

During the last few years the term cyber risk definitely shows signs of growing in popularity. This is somehow connected with the process of digitisation and interconnectivity of systems in organisations. While business organisations rely actively on IT, and their employees use more and more applications on their laptops and mobile devices for the purpose of their work, this leads to the generation of enormous amounts of available data and information. Protecting this data, applications and systems from different forms of cyber security risk via the Internet like a denial of service (DoS) attack or data theft is a task that should not be underestimated.

Cyber security risk is defined as a risk to information and technology assets that have consequences which affect the availability, integrity and confidentiality of information or information systems, while the sources of cyber security risk can be organised into four classes: actions by people, systems and technology failures, failed internal processes and external events [2]. According to another definition cyber risk is defined as the chance of damage or loss from an electronic exposure that can result in a negative impact for a business organisation like disclosure, destruction or theft of data or unavailability of applications, systems or networks [3]. Cyber risk is not a specific risk but it is presented as a group of risks which differ in technology, means and attack vectors and have a potentially great impact on the target – from legal liability and computer security breaches to privacy breaches and theft of confidential data [1].

Managing cyber security risk means that business organisations need to pay more attention to cyber security which, by premise, is an extension of established IT security. Traditionally, boards have tended to believe that questions concerning the security and integrity of corporate data is in the area of their IT staff. However dealing with cyber risk is changing the focus from a purely technical issue to a serious business risk and that's why it should involve the attention of top executives and the board of directors [4].

One major reason for this new attitude to cyber risk is liability. Nowadays there are a lot of demanding regulations in North America, Europe and United Kingdom concerning the management of personal data – for example the EU General Data Protection Regulation, the Sarbanes-Oxley Act and The Health Information Technology for Economic and Clinical Health Act. Their fines and penalties are very rigorous when it comes to stolen data and at the same time there is the possibility of criminal investigation, prosecution or legal claims from customers or employees. The above mentioned facts lead to the conclusion that compliance with these laws should be a huge priority for the governing body of every business organisation.

The growing numbers of targeted cyber-attacks which are increasing in their frequency and scale points out another reason why managing cyber security risk should no longer be seen only as an IT issue. According to the survey "2013 Cost of Data Breach

Study", conducted by Ponemon Institute, more business organisations are looking to involve not only the Chief Information Officer but also other corporate leaders in the management of cyber risk while the average cost of a data breach is 188$ for each lost or stolen record [7]. The latter research paper also noted the following considerable findings:

- ✓ Security exploits and data breaches lead to a financial loss of, on average, 9.4$ million for a company;
- ✓ Concerns about cyber risks rank as high as or higher than risks from natural disasters – 76% of respondents share this opinion;
- ✓ Increasing interest in cyber security insurance – 31% of the surveyed organisations already have a policy and 39% of them plan to purchase one in the near future.

All this shows that cyber risk has evolved and essentially it could affect any business that utilizes or uses technology in its day to day operations, or handles or stores personal and confidential data. Many recent surveys (by Kaspersky Labs, Harvard Business Review, Deloitte, Marsh etc.) have come to the conclusion that these risks are increasing in scope, frequency and severity and need to be managed in an appropriate manner, but actually nearly two thirds of incidents are caused by issues (employee negligence, system glitches or malicious attacks) which should be controlled at reasonable levels by business organisations [6].

3. Cyber Risk as a Business Continuity Planning Issue

A business continuity plan should be created, tested and implemented to ensure the recovery of mission-critical business operations in a timely manner. During the whole business continuity process, information security needs to be woven as an integral part of it. Business continuity planning needs to address information security requirements and must conform to them as every other process in the organization should.

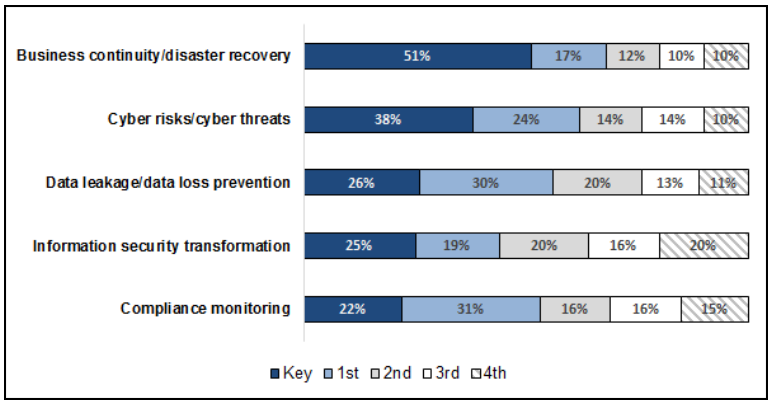


Fig. 1. "Top priority" information security areas over the coming 12 months.

It is not a surprise that business continuity/disaster recovery is indicated as a key information security area for the next 12 months with 51% of respondents marking this answer (see Fig. 1) [8]. The second top priority area is cyber risks/cyber threats. This shows that returning the organisation to "business as usual operations" as soon as possible after disruptive events is a goal that most businesses have already understood, and they would like to implement this kind of policy. Meanwhile cyber threats have become a real business risk and need to be

managed appropriately not only in the area of information security but also as a part of business continuity planning.

Looking at the results in the global technological landscape in 2013 compared to 2012 (see Fig. 2) we can easily see that cyber-attacks and incidents of data fraud/theft register a rise in both the likelihood of occurrence and the impact of the event. This is another confirmation that cyber risks do exist and that cyber security should be treated as a board-level issue within business organisations. At the same time, like critical systems failure, cyber-attacks depict a possible scenario of disruption of normal business operations. Thus, during the risk assessment phase of the business continuity planning process, sufficient time should also be devoted to cyber risk assessment.

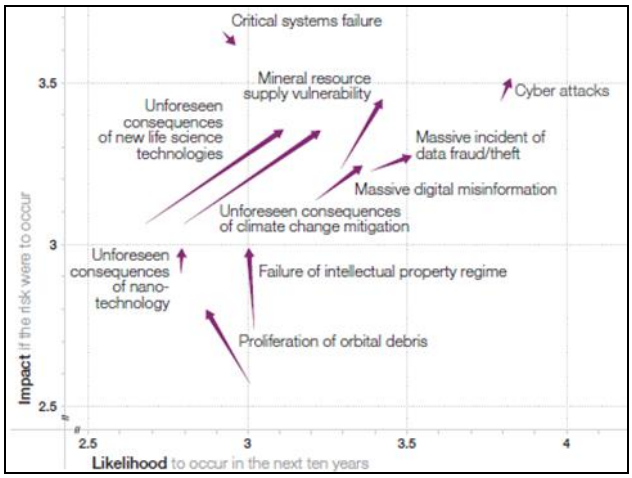


Fig. 2. Global Technological Risks Landscape 2013 versus 2012 [5].

This requires a proactive approach to identify and deal with cyber challenges. It means that business organisations should analyse and understand the connection between their business operations and mission requirements on one side, and the cyber domain on the other side. It is important to identify essential cyber assets and what key business functions and mission capabilities they support. Doing so guarantees the achievement of cyber resilience and helps in reckoning the impact of loss due to a certain cyber event.

Including cyber risk when planning business continuity helps with testing the performance and capabilities of cyber assets and also supposes that there can be weaknesses found in cyber-related operations which could be referred to as areas for improvement. The business continuity plan is not complete if it doesn't take into account the need to maintain the availability of the mission-critical cyber assets of the organisation. There is a bilateral effect – a plan as to how to keep a business "up and running" in a case of cyber-attack but also to improve cyber security through the business continuity management lifecycle.

4. Conclusion

The consideration of the above leads us to the conclusion that as more and more activities in and between business organisations and their customers and business partners move through interconnected systems and the Internet, cyber security risks will multiply in any sense. Business organisations that have implemented a cyber security policy and have considered the potential disruption of critical business functions as a result of cyber-attacks are prepared to mitigate and minimise the damage from a security breach. Their business continuity plan will keep them operating and doing business, while other competitors which are less prepared might fail, thus having an extensive impact on their reputation, revenues, employees' morale and future business opportunities.

References

1. Barzilay M. A Simple Definition of Cybersecurity, 2013, (http://www.isaca.org/Knowledge-Center/Blog/Lists/Posts/Post.aspx?List=ef7cbc6d-9997-4b62-96a4-a36fb7e171af&ID=296)
2. Cebula, J., L. Young. A Taxonomy of Operational Security Risks, Carnegie Mellon, 2010.
3. Cyber Risks: How to protect your business in the Digital Age, 2011, (https://www.immersionltd.com/Immersion/documents/CyberRisks-Business-Insurance-White-Paper_Jan2011.pdf)
4. Dunbar, T. The First Step to Managing Cyber-Risk, Risk Management (New York), 59, 2012, N 8, 20-24 p.
5. Global Risks 2013 Eight Edition, 2013, (http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2013.pdf)
6. Lloyd's Risk Index 2013, 2013, (http://www.lloyds.com/~/media/Files/News%20and%20Insight/Risk%20Insight/Risk%20Index%202013/Report/Lloyds%20Risk%20Index%202013report100713.pdf)
7. Managing Cyber Security as a Business Risk: Cyber Insurance in the Digital Age, 2013, (http://www.experian.com/innovation/business-resources/ponemon-study-managing-cyber-security-as-business-risk.jsp)
8. Under Cyber Attack: EY's Global Information Security Survey 2013, 2013, (http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2013.pdf)

Electronic verification of public projects

Katia Kirilova[1]

[1] UNWE, 1700 Sofia, Bulgaria
kstrahilova@abv.bg

Abstract. The report examines the current methods for verification of public expenditure projects. It presents the characteristics of public sector projects. They are given basic methodological formulations for the verification as a business process. Below are the characteristics of the process and its characteristics. The main results are related to the electronic verification of accounting documents scanned.

Keywords. Information Technologies, public projects, electronic verification.

1. Introduction

The project BG051PO001 - 3.3.07-002 "Student Practices" (the project ends on October 31, 2014) is implemented by the University of National and World Economy in partnership with the Ministry of Education and is co-financed by the European Social Fund [2]. Its objectives include improving the quality level of education by providing opportunities to gain experience and improve the practical skills of students of University of National and World Economy in accordance with the needs of the labor market.

2. Approach for uploading scanned documents in the system

The project is based on the approach to scan and upload in web based platform of all contracts and supporting documents. Project is based on the principles of the computerization of the management and verification of funds. Some principles of this process are described by Kirilov [1]. Basically similar approaches to scan and upload scanned documents signed in the systems are already common in practice and demonstrate their effectiveness. In particular it is (see Figure 1 and Figure 2):

- Contract between university and student;
- Contract between university and employer organization;
- Contract between university and the Ministry of Education;
- Contract between university and expert;
- List of insurance;
- Order to determine the professional staff;
- Contract between university and mentor.

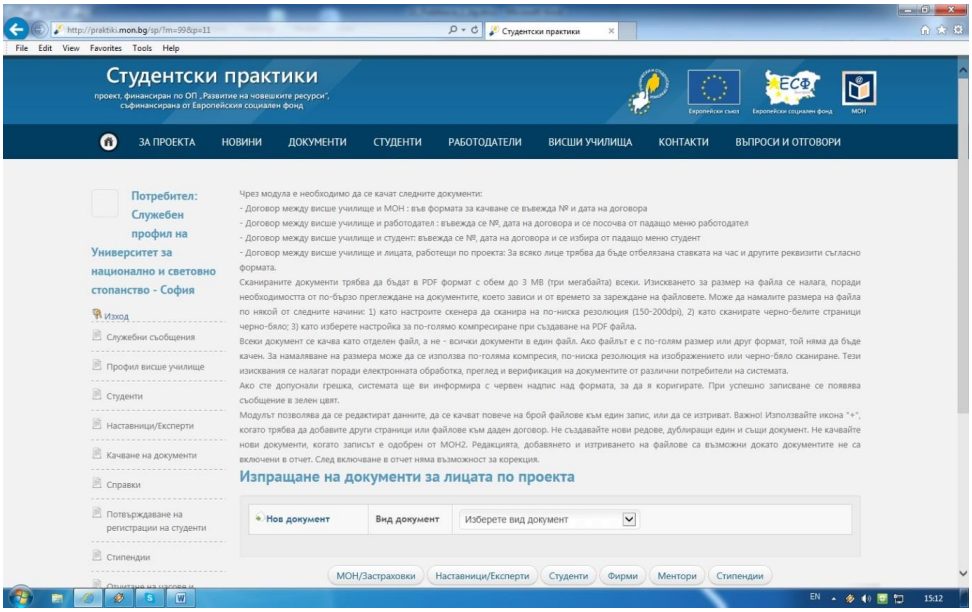


Figure 1. Sending documents for individuals within the project

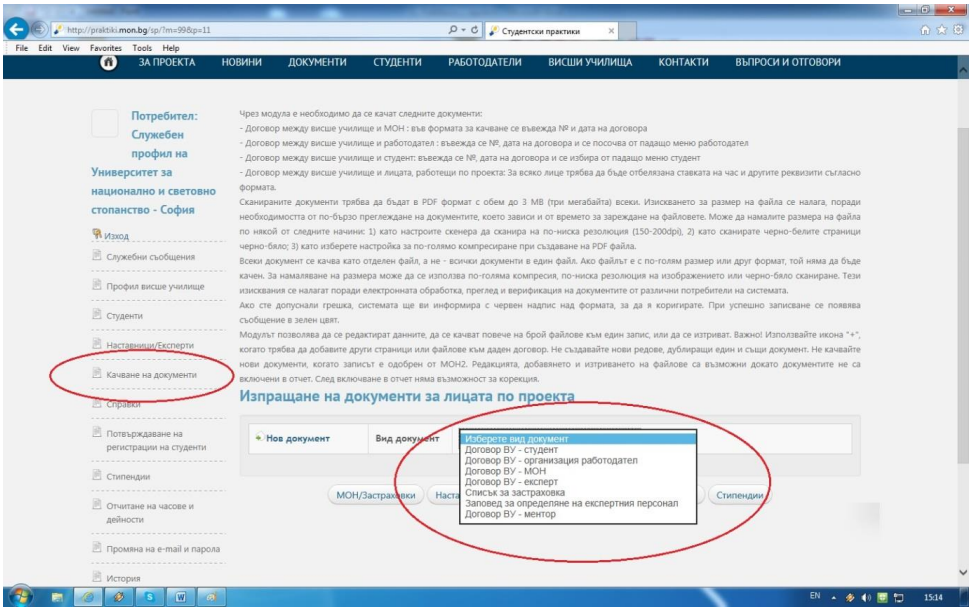


Figure 2. Document types

At Figure 3 is shown an option of a verified document.

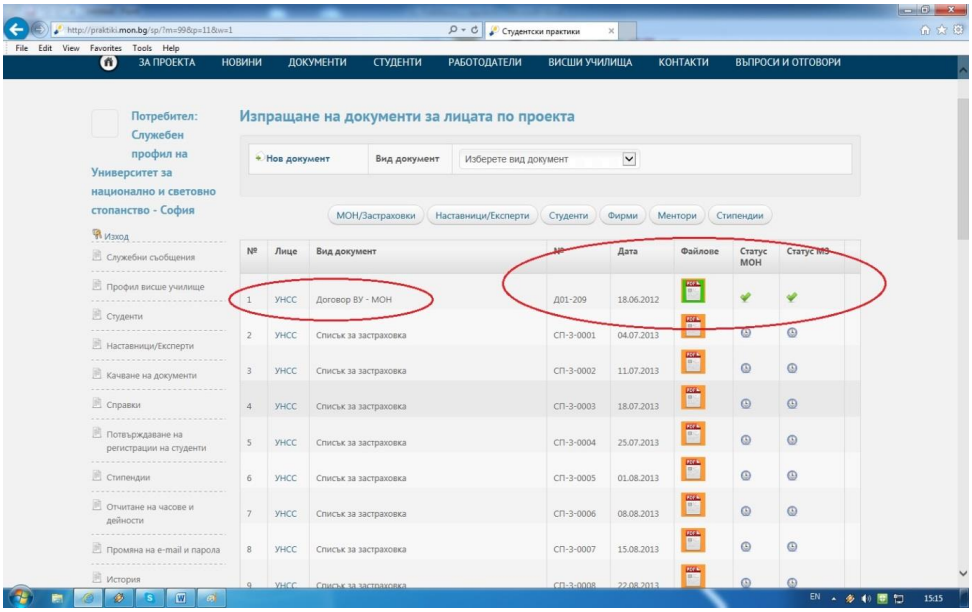


Figure 3. Scanned and verified contracts between the University and MON

The main levels of verification of these documents are:

- Document uploaded from high school;
- Status MON;
- Status MZ.

At Figure 4 is shown an option of verifying the cost for scholarships to students.

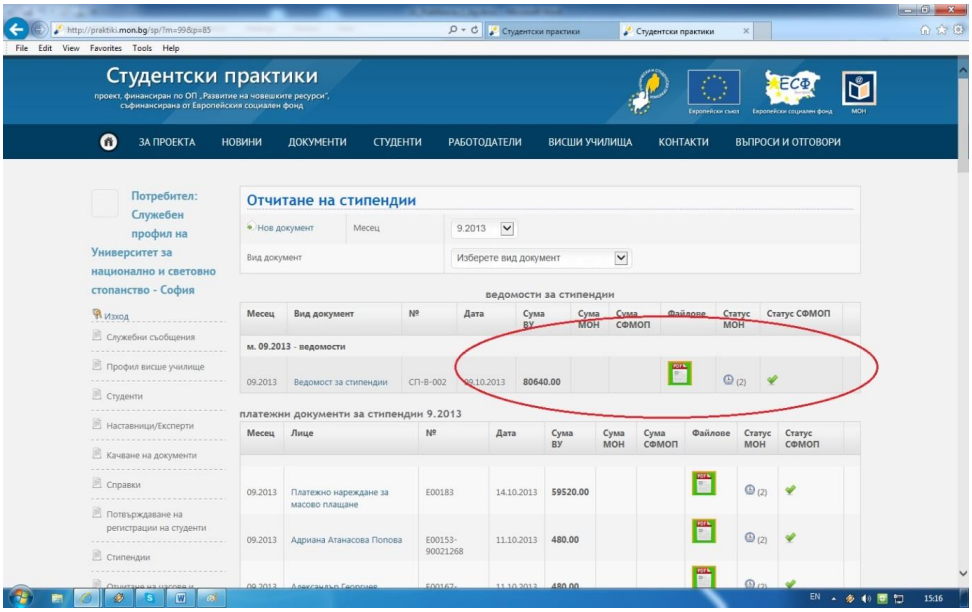


Figure 4. Electronic verified expenditure

The main levels of verification of these costs are:

- Type of expenditure uploaded from high school;
- Status MON;
- Status SFMON.

In the example case, this cost is successfully verified on three levels.
At Figure 5 is presented an opportunity to verify the total cost.

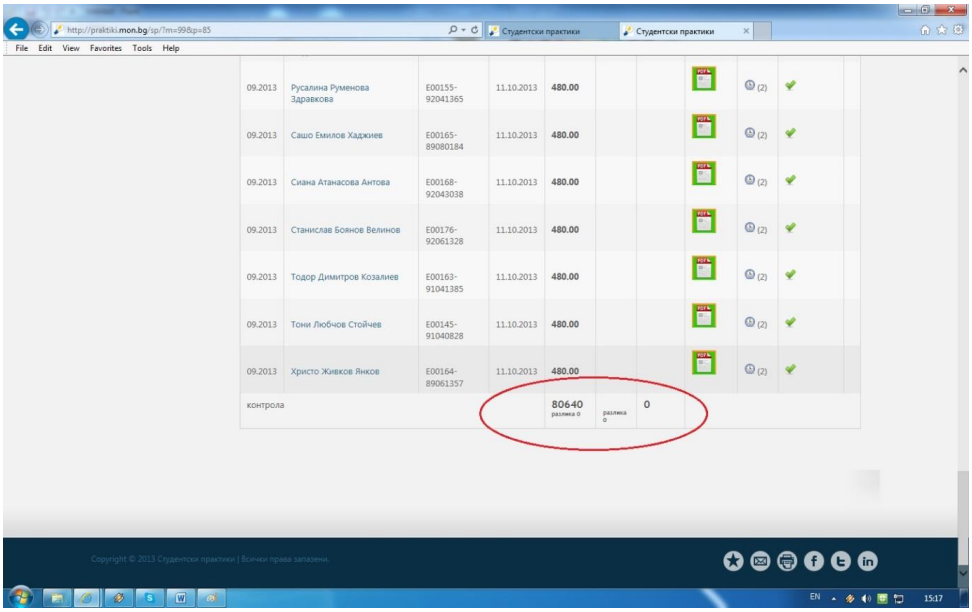


Figure 5. Verification of general expenses

3. Conclusion

The main purpose of University of National and World Economy with this project is to provide students with knowledge, skills and competencies as a result of training and work in a real practical environment that creates opportunities for better realization in accordance with the requirements of the labor market. Already a part of students receives concrete proposals for employment. Electronic reporting within project BG051PO001 - 3.3.07-002 "Student Practices" makes possible the application of a new approach for verifying the costs of public projects.

References

1. Kirilov, R., "Building e-municipality, problems, approaches and solutions", “Trakia Journal of Sciences”, Vol. 11, Suppl. 1, 2013, Trakia University, ISSN 1313-7069, 2013.
2. Official website of project “Student practices”, http://praktiki.mon.bg/sp, 02.12.2013.

Role of practical training in public administration to increase student's knowledge and competencies

Katia Kirilova[1]

[1] UNWE, 1700 Sofia, Bulgaria
kstrahilova@abv.bg

Abstract. Conducting practical training is very important for students. It carries the combination of theory and practice. This is the key to easy adaptation of graduates in the labor market. Practical training is important for students of public administration. It allows upgrading their knowledge, skills and competencies in line with the labor market. The main objectives of the training are: to help improve the quality of education to facilitate the transition from education institutions to the workplace to enhance the successful implementation of the youth labor market, to help build strong partnerships between educational institutions and business, to strengthen the incentives for students to participate in additional practical training in a real work environment. In conclusion, formulate conclusions.

Keywords. Information technologies, public projects, electronic verification.

1. Introduction

The project BG051PO001 - 3.3.07-002 "Student Practices" (the project ends on October 31, 2014) is implemented by the University of National and World Economy in partnership with the Ministry of Education and is co-financed by the European Social Fund [3]. Its objectives include improving the quality level of education by providing opportunities to gain experience and improve the practical skills of students of University of National and World Economy in accordance with the needs of the labor market.

2. Contracts – execution and administration

For the period from July 1, 2013 to January 31, 2014, the number of contracts for student practice in the University amounted to 1930 units (Table 1).

Table 1. Number of contracts

Registered students within the information system of the project	Number of practices confirmed by employers	Number of practices confirmed by students	Number of contracts
6135	3128	2960	1930

A graphical representation of the registered student is given at Figure. 1

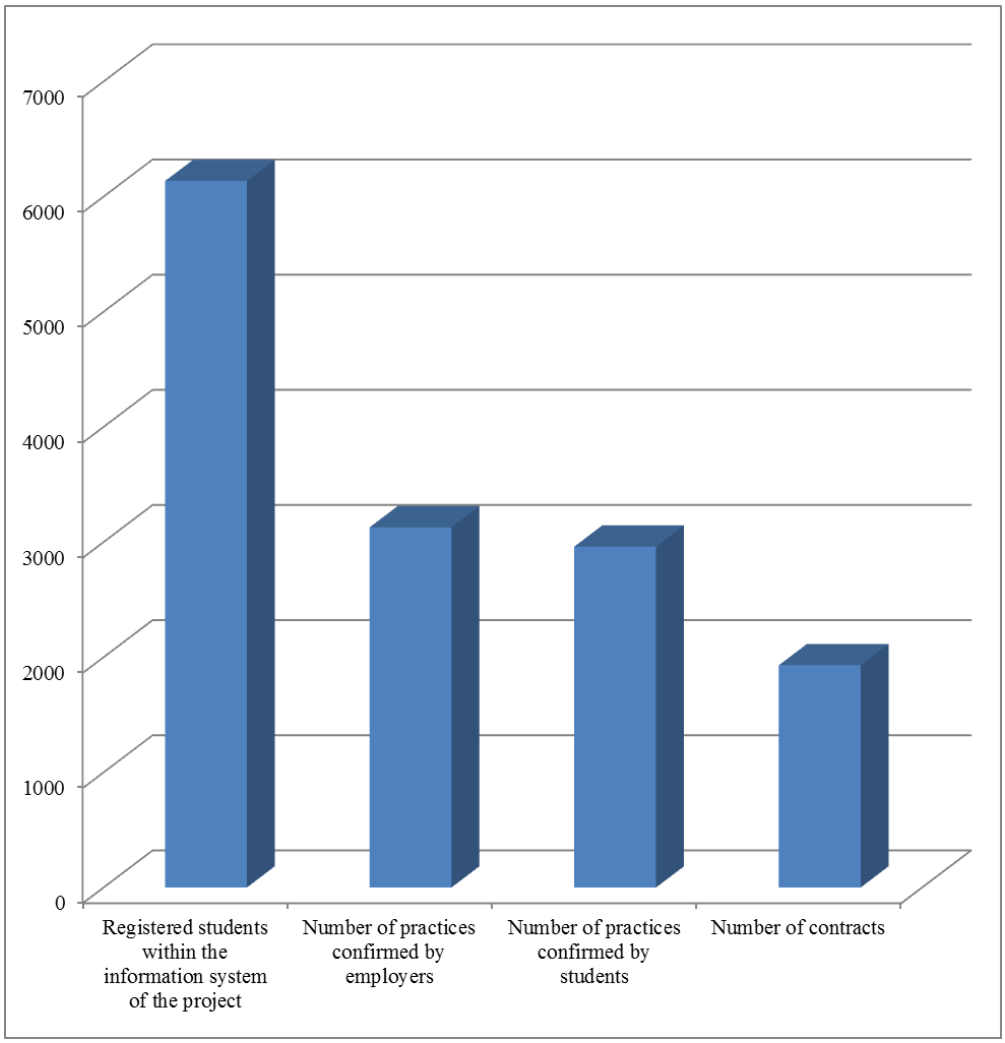


Figure 1. Number of registered students and contracts

3. Participation of students of "Public Administration"

Computerization of the process of application and cost accounting fits perfectly into the current trends of e-governance. Similar principles are described by Kirillov [1].

To participate in the practice can apply all the students registered in the register of current and discontinued undergraduate and graduate students supported by the Ministry of Education. After a survey we found that students of University of National and World Economy involved with the project are mainly students bachelor's degree.

University of National and World Economy in Sofia participates in the project with 6135 registered students and so far 1930 of them are contracted to conduct practical training. Distribution of contracts by departments is shown at Table 2.

Table 2. Distribution of contracts by departments

General Economics	283
Business Faculty	384
Faculty "Management and Administration"	291
Financial - Accounting Faculty	345
Faculty "Economics of Infrastructure"	217
Faculty "International Economics and Politics"	144
Faculty of Law	226
affiliate "Economics and Management"	7
Faculty "Applied Informatics and Statistics"	62

A graphical representation of the distribution by departments is given at Figure 2.

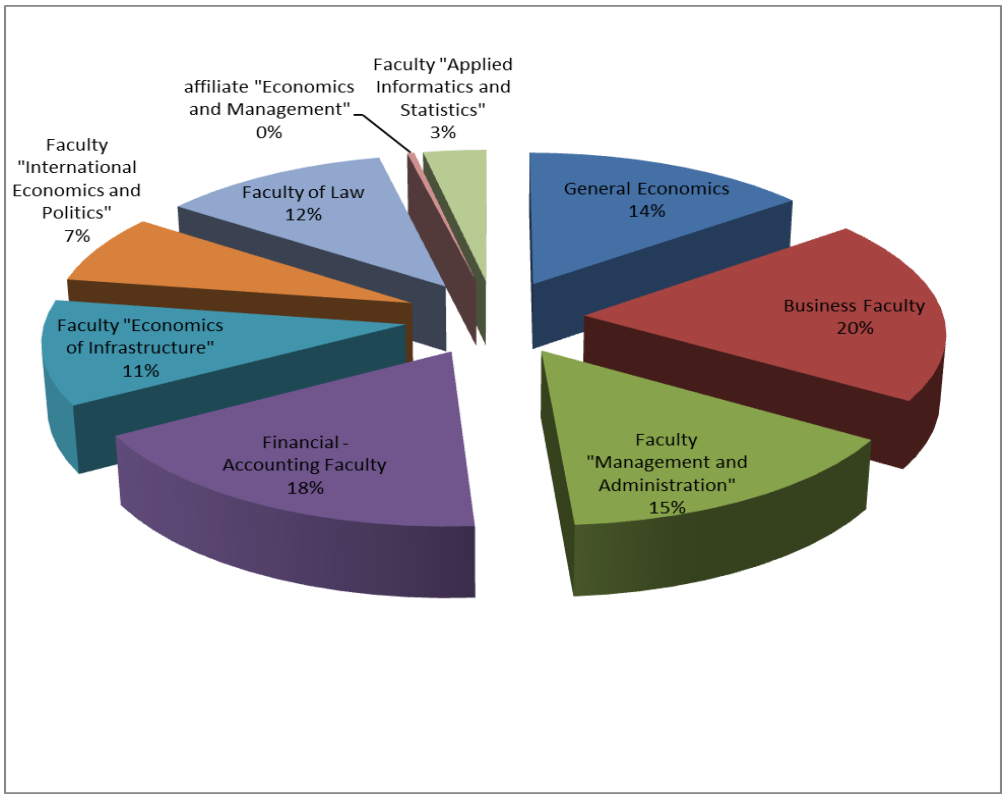


Figure 2. Distribution by departments

Application, approval and reporting of practices / internships takes place in a specially constructed for this purpose online system (praktiki.mon.bg). Architecturally this system represents classic three-tier architecture. Database study of such a system gives rise to the formation of different nature and structure records. Detailed description of databases of such systems is done in a study by Milev [2]. He points out the special importance of the process

of dynamic database structure of a similar system. This creates opportunities to facilitate the students with the system. Every student can apply to more than one employer and can be selected by more than one. Students sign a contract for practical training in the University of National and World Economy in 12 days of confirmation of an offer from an employer.

As seen from the data presented in Table 2 and Figure 2 the participation of faculty "Management and Administration" – the composition of which is the specialty "Public Administration" is significant - about 15%.

Practices of this project are not mandatory for students, despite the interest in them is very large. Within the project, the students undergo practical training in a real working environment for 240 hours. After the Successful completion of the internship students receive a scholarship of 480 lv. Until 31 October 2014, the expectations are that the number of contracts with students of the University is going to reach 3100.

Increasing interest in the practices of students from University of National and World Economy account among employers - our students are preferred by banks, consulting firms in finance , insurance companies, manufacturing companies, accounting firms, law firms, municipalities, government agencies, non-governmental organizations, museums, real estate agencies, publishers and others. By the time University of National and World Economy has contracted with over 500 employer organizations and over 1000 mentors from them that assist students in their practical training.

4. Conclusion

The main purpose of University of National and World Economy with this project is to provide students with knowledge, skills and competencies as a result of training and work in a real practical environment that creates opportunities for better realization in accordance with the requirements of the labor market. Already part of students receives concrete proposals for employment. Implementation of the project BG051PO001 - 3.3.07-002 "Student Practices" is in support of the mission of the University of National and World Economy to provide their students with knowledge and skills so that they creatively solve the problems of the present and the future in the fields of economics, management and administration, law and politics.

References

1. Kirilov, R., "E-Government in the public sector - challenges and paradoxes", International conference on application of information and communication technology and statistics in economy and education, University of National and World Economy, ISBN 978-954-92247-4-0, 2013.
2. Milev, P., "Opportunities for building computer information systems in municipalities using dynamic document management", E-Journal "Dialog", Issue 2, 2013, Tsenov Academy, ISSN 1311-9206, 2013.
3. Official website of project “Student practices”, http://praktiki.mon.bg/sp, 02.12.2013.

FORECAST OF FINANCIAL CONDITION BASED ON ITS OWN SAMPLE

Lewandowski Vladislav

Academy of Economic Studies of Moldova
Depechevl70@mail.ru

Abstract. In theses is treated a new method of forecasting the financial condition of the company. The algorithm is interesting because in practice shows the high accuracy of the forecast. This is achieved by eliminating the factors of reduced quality prediction, which characteristic of the traditional forecasting methods.

Keywords: forecast financial position, financial performance, financial ratios.

One of the advantages of template method to the classical methods of prediction, it is the fact that classical methods have been developed based on a sample of different types of businesses, but the template method of forecast in addition to the data of other companies, the system can get forecast, exploring its own sample. However, at the beginning of the operation of the business, such opportunities are not available, because not yet formed its own sample with accounting data. Some people believe that there is no difference, what kind of samples in using, their own samples or samples of other companies. Nevertheless, following the logic of common sense when using the sample of other enterprises (mixed sample) template method is no longer taken into account type, size businesses, activities, economic conditions and other features of the functioning of the company. This omission can not but have a negative impact on the quality of forecasts.

In this regard, there is a need to verify the accuracy of the template method forecasts based on own sample and a sample of the generated data from different companies.

To test it was decided to form a sample of data from different companies (Safin-Grup, DigiAger, GekaTex, CarComVit, Bilgi-Com). When this sample was obtained, it was necessary to accommodate the data so that the reporting periods within each company located consecutively, and also the sample of companies have to be placed one after the other.

Under this condition, the algorithm is guaranteed the correct work of template method of the forecast. (LINK to article Lewandowski). To fulfill this requirement, in the field of date, have entered values with increasing periods of the quarter and the year. For example:

date
31.03.1995
30.06.1995
30.09.1995
.....
31.12.2012

Thus, initially having a five-year quarterly sample of five companies, was derived the database which consist from 100 lines. One line contains the data for one quarter.

Before we consider how to perform the verification of prediction accuracy, recall the essence of the template method of forecast financial condition. The financial condition of the company presents itself continuously changeable the value, which can be characterized by a set of values of financial performance for the period. If the company is solvent, its state is characterized by a specific set of financial ratios, characteristic for successful financial

situation. When an organization is in crisis, its state can be characterized by sets of indicators specific to the crisis. When a company is on the eve of the crisis, respectively, such financial situation can be characterized by a set of indicators specific to the pre-crisis situations.

The essence of a generic prediction method is that the system searches in the database with sets of indicators for various periods, such a set of indicators that would be similar in the values of the current reporting period. If such a set is found, and this was found to be characteristic of a set of pre-crisis situation, then the system should warn the user about the possible development of the crisis. Pre-user should be noted in the database all the sets corresponding to the periods of crisis as a "crisis." And all the periods that preceded the crisis, respectively, should be labeled as a "pre-crisis".

In general, the template method can predict any of the financial condition, it should not necessarily be a "crisis." To form a prediction, it is enough to found in the database a set of values of indicators that would be similar to the set of the current reporting period. Subsequent set, after found set in the database will be a set of prognostic value.

The significance of the template method of forecast has been shown in the prior «Template method of forecasting the financial condition» [2].

It should be noted that in obtaining results (Table 2), the calculations are not performed in a mixed databases, but in separate databases. Each company counted prediction accuracy only according to their own selection. Now it is necessary to perform the same calculation of accuracy of the prediction, but based on the newly formed sample of data from different companies. The algorithm of the calculation is identical to the algorithm calculating of predictive accuracy forecast on their own sample, which was discussed in detail in the second chapter, the difference in the treated samples (on their own sample or are of different types).

But still, to get a forecast on the basis of samples of different companies still need to make some changes. Yet from the results obtained by polytypic (mixed) sample subtract the number of combinations which made up from periods of the own sample. In this case, the results will be only a combination of different sets of indicators of enterprises. What is required.

Table 1

The final results of calculations of the forecast accuracy based on of the polytypic samples

16%	of different types of sample R	Own sample S	different types of sample without own sample W= R - S
YYY	581	53	548
YYN	125	15	110
YNY	2394	94	2300
YNN	669	27	642

For example, 16% error comparison were obtained the following results shown in Table 1 in the third field "private sample." YYY designation means that the values of the indicators coincided, dynamics and prediction. Y-from the English «YES». The first "Y" means that coincided indicators, the second "Y" - the dynamics coincided, and the third "Y" means that coincide forecasts. «N», respectively, taken from the English «NO», that is no

coincidence. If the second N, then did not coincide the dynamics and on the third N - did not coincide predictions.

Then, using the data in Table 1 from the fourth field of W, by the formulas (1), (2) and (3) were calculated with the accuracy of forecasts with the coincidence TYYY dynamics and coincidence only indicators TYNY.

TY:=(YYY+YNY)*100/(YYY+YYN+YNY+YNN)=(548+2300)*100/3600=79,11% (1)

TYYY:=YYY*100/(YYY+YYN)=548*100/(548+110)=83,28% (2)

TYNY:=YNY*100/(YNY+YNN)=2300*100/(2300+642)=78,18% (3)

The results were listed in Table 3 in field "error 16%". By A similar scheme was calculated accuracy and for the another errors of the comparisons, all the results rendered in table3.

Table 2

Average value of the precision with the error of comparison from 5% to 25%

		25%	20%	16%	10%	5%	Average, %
SR	**TY**	80,21	75,95	75,03	69,6	70,02	**74,162**
SR	**TYYY**	91,88	83,75	74,22	70,05	69,31	**77,842**
SR	**TYNY**	76,83	73,09	76,1	67,62	64,47	**71,622**
The difference: TYYY – TYNY		**15,05**	**10,66**	**-1,88**	**2,43**	**4,84**	**6,22**

Table 3

The average value of accuracy with an error of comparison from 5% to 25% of the mixed sample (b-blend)

		25%	20%	16%	10%	5%	Average, %
SR	**TYb**	93,68	89,1	79,11	53,39	28,37	**68,73**
SR	**TYYYb**	94,21	90,58	83,28%	56,82	38,64	**72,71**
SR	**TYNYb**	94,5	88,77	78,18	52,09	25,7	**67,85**
The difference: TYYYb – TYNYb		**-0,29**	**1,81**	**5,1**	**4,73**	**12,94**	**4,86**

Comparing the average values of Table 2 and Table 3 can be obtained a table of the difference # 4, which is obtained from the difference between the values in Tables 2 and 3.

Table 4

The difference between the accuracy of the prediction between the values obtained on its own and mixed (multi-type) samples

TY-TYb	5,433%
TYYY-TYYYb	5,132%
TYNY-TYNYb	3,772%
Среднее	**4,8%**

The average value of accuracy of the overall forecast of template method on its own sample turns 74,162%, but in a mixed sample of 68.73%, a decrease of 5.433%. Forecast accuracy, taking into account of the retrospective dynamics, based on his own sample is higher about on 5,132%, than a similar outlook on a sample of different types of businesses. Forecast without consideration of the dynamics indicators on their own sample is higher on 3,772% than the forecast based on a sample of different types of businesses.

Thus, the forecast of the financial situation obtained by the template method and formed on their own of accounting data sample is higher by about ≈ 5%, then the similar forecast based on a sample of other companies.

Reference

1. Bogdanov T.K. The use of multivariate statistical analysis to predict the probability of bankruptcy Russian companies / / Proceedings of the IX International Conference "Application of multivariate statistical analysis in economics and quality assessment" (24-26 August 2010). - Higher School of Economics, 2010. - P. 64-67.
2. Lewandowski V.I. "Template method of forecasting the financial condition" United Kingdom, London, International Academy of Higher Education September 5, 2013 http://gisap.eu/node/32130#comment-32397

Information Architecture Design of a Business Information System for Performance Management of Business Applications

Alexandrina Murdjeva[1], Veska Mihova[2]

[1] UNWE, 1700 Sofia, Bulgaria
amurdjeva@unwe.bg

[2] UNWE, 1700 Sofia, Bulgaria
vmihova@unwe.bg

Absract. With the development of information systems and the growth of managed and stored data arises a need for monitoring and management systems. The companies are increasingly interested in the performance of the systems which manage their business. This report is a result of current and previous research for systems' performance and more specifically the focus is on the database productivity in a system. The database is considered as a part of the entire business intelligent system. This report will propose information architecture of a business information system for performance management of a database. The data sources for business information system for performance management of business applications will be described as long as the approaches for collecting and processing data in the system.

Keywords. Database, database monitoring, database administration, database performance, performance forecast.

1. Introduction

The problem of business information systems' monitoring and optimization has been a current topic recently because the need for a solution is recognized in more and more business areas.

Performance consulting and finding a solution to a performance issue is a part of database administrator's work.

Previous research papers present a concept for database performance management. The entire idea is focused on early recognition of problems with database performance management and their solution. This report proposes an information architecture of a business information system for database performance management. There are some specifics of the data sources of the system described, as well as methods and tools for collecting and processing the system's data.

2. Information architecture

We will start with the entire figure, figure 1 below visualizes the information architecture of the system for database performance management. Then we will focus on each of the layers to outline their main features.

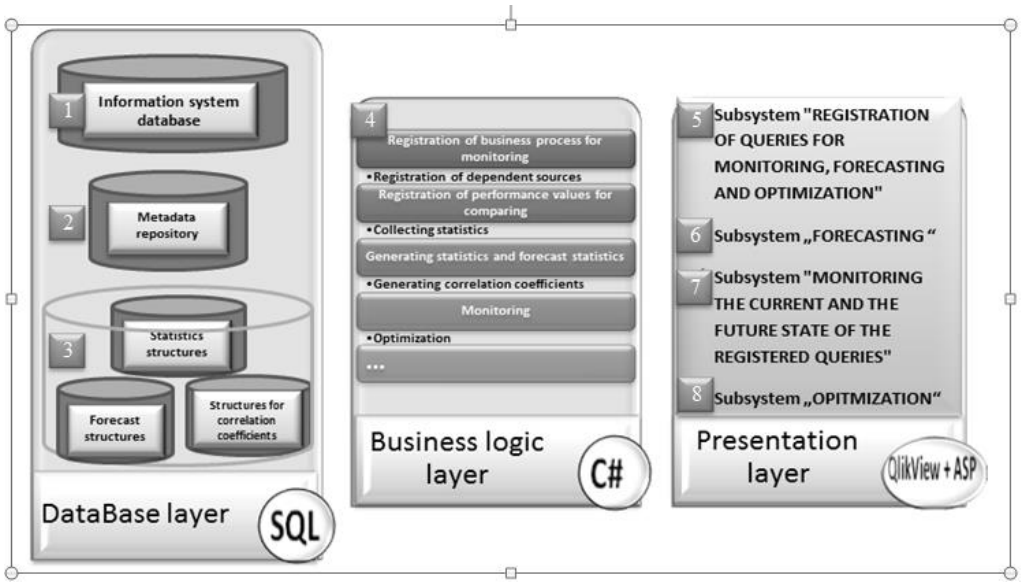


Fig. 1. Information architecture design of a business information system for performance management of the business applications.

a. Database layer

The data sources in the proposed architecture are the database of the business information system, the database that stores metadata for the management process of generating the structure of the data warehouse and the structures that store performance data – statistics, forecast statistics and correlation coefficients.

i. Information system database

The first data source that is the system's main data source is the database or databases of the business information system that will be monitored and examined. Each DBMS can provide statistics about the work of the system at a certain moment, but most of the values are momentary, without accumulation, which means that we can't follow the degree of change in the execution time for a period of time. For that reason we propose structures for holding momentary states of the database, so that we can store and analyze more than one momentary state. The main issue that arises here is that because of the different structures in which the different servers store data about the work of the business system, we need to have different structures for each of them, too. This fact leads to the idea of storing metadata in the database performance management system, which is based on the current information communication architecture.

ii. Metadata repository

The metadata is stored with the set purpose for automatic generation of data warehouse according to the specifics of the database servers. This metadata can indicate

what the corresponding database's name is, what DBMS it works with, and the path to the database, according to how long it takes to make a forecast and other relevant data. The idea is to generate a data warehouse with a structure that is appropriate to a different DBMS. The data warehouse is designed to store historical information that is extracted from the system tables and tables with statistics of a particular DBMS so its data source is potentially one and is homogeneous. Therefore, it's reasonable to propose storage with an analogous structure to the system tables' structure, equipped with an indicator of time. In such a structure it is possible to provide procedures for generating these structure-based meta descriptions.

iii. Database for performance management of the business applications

1. Statistics structures

The statistics structures are loaded by ETL processes. The stored data in the specific structures is statistics for the execution time of queries, for data volumes in different database objects and other necessary relevant statistics for generating forecast values.

ETL tools, stored procedures, offline processing or other appropriate tools provide generation capability of statistical values at a precise time, depending on how dynamic the process of system operation is.

Fig. 2. Statistics data structures for data sources, on which the query depends – SQL Server.

2. Forecast structures

After collecting enough historical information about the operation of the system, predicted values are generated based on the gathered statistics periodically. The generation of

these forecasted statistics can be done using statistical packages, Data Mining, stored procedures, MS Excel and many other instruments.

In the proposed architecture forecast structures with forecasted data are provided. This is still storage structure, since this is historical data, accumulated values for different periods, with the only specific that these periods are in the future.

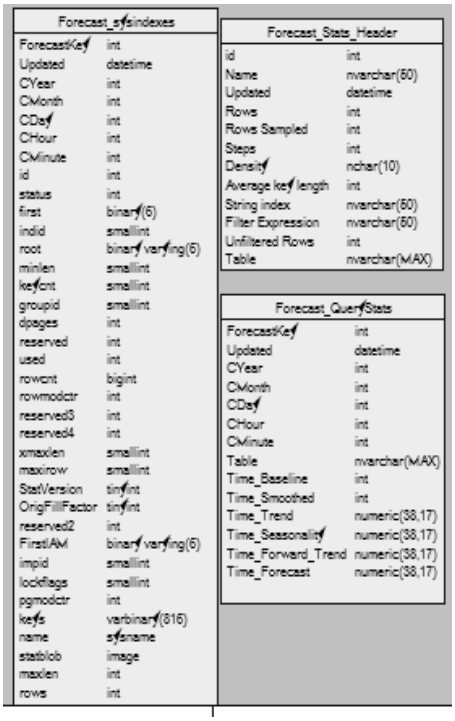


Fig. 3. Forecasted data structures for data sources, on which the queries depend – SQL Server.

3. Structures for correlation coefficients

Apart from structures with forecasted information about data values that are stored in the system and the execution time of queries, the architecture makes the collection of correlation statistics between data volumes and the execution time of queries for a certain period of time possible.

The correlation coefficient between the two variables is calculated and stored in other specific structures for correlation coefficients.

As the survey shows, at different moments of time one query is affected by different resources. The idea is that the system can accumulate historical information about these dependencies and, within a particular time, can record their correlational value.

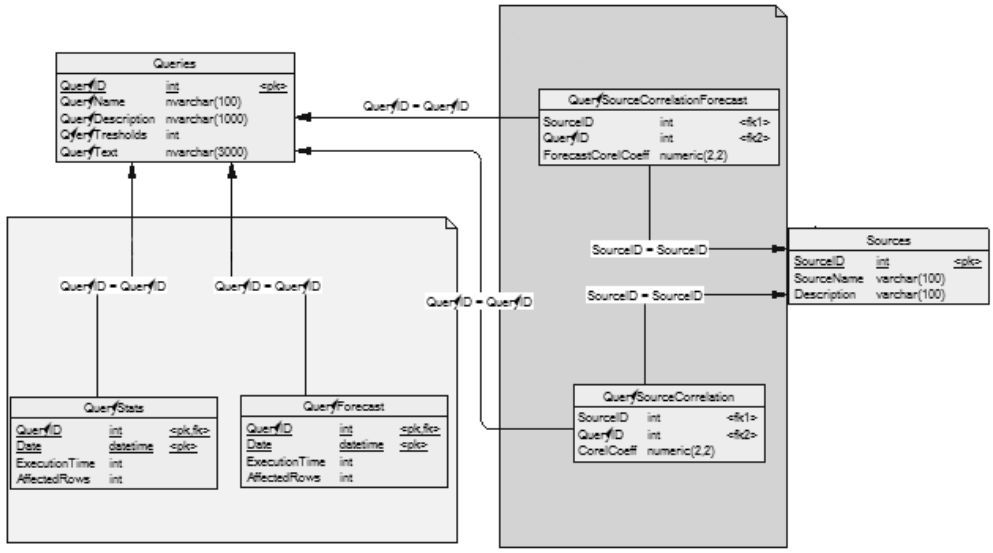


Fig. 4. Structures for describing queries, sources and correlation coefficients.

b. Business logic layer

This layer is supposed to assure data registration, processing and execution of steps included in all subsystems. The business logic is developed using the C# programming language.

The business logic layer includes functionality such as:

- Collection and management of data for generating a specific data warehouse;
- Registration of a server, a database and a way of authentication;
- Registration of a business problem for monitoring;
- Registration of data source, on which the registered query depends;
- Registration of values for comparing performance;
- Other.

c.Presentation layer

This layer includes presentation of the results from the four subsystems:

- Subsystem "Registration of queries for monitoring, forecasting and optimization"
- Subsystem „Forecasting “
- Subsystem "Monitoring the current and the future state of the registered queries"
- Subsystem „Opitmization“

This part of the system can be developed using different tools. For the current research we have chosen ASP.Net application.

3. Conclusion

In summary, in this paper an information architecture design of a business information system for performance management of the business applications is presented.

It is important to have a tool for database performance management in order to identify potential risk areas for improvement. We have developed a prototype system for database performance management based on the proposed architecture. The architecture, presented in this report, gives database administrators a new level of database performance management and monitoring. The solution gives the database administrators enough time for reaction to an occurring problem with the application's performance and respectively the database's performance.

References

1. Eric Hu, sap-application-development-performance-control http://www.sap-perf.ca/sap-application-development-performance-control/
2. Murdjeva A., Mihova V., Architecture of a business intelligent system for monitoring the present and the future state of a database, International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE 2012 UNWE, Sofia , Bulgaria, 5-6 October 2012
3. Yuhanna, Noel, Penn, Jonathan и Smillie, Katie . Forrester Research:Research: The Forrester Wave™: Enterprise Database Auditing And Real-Time Protection, Q4 2007. [Онлайн] October 2007 г. http://www.forrester.com/The+Forrester+Wave+Enterprise+Database+Auditing+And+RealTime+Protection+Q4+2007/fulltext/-/E-RES41970?objectid=RES41970.
4. Oracle9i Database Performance Tuning Guide and Reference Release 2. [Онлайн] http://docs.oracle.com/cd/B10501_01/server.920/a96533/stats.htm.
5. Database monitoring best practices: Using DAM tools. [Онлайн] http://searchsecurity.techtarget.com/tip/Database-monitoring-best-practices-Using-DAM-tools.
6. Proctor, Paul E., Litan, Avivah и Nicolett, Mark. Pattern Discovery With Security Monitoring and Fraud Detection Technologies. [Онлайн] 2 September 2009 г. http://www.gartner.com/DisplayDocument?id=1160514.

Latest Trends in Business Intelligence System Development

Dorina Kabakchieva, Kamelia Stefanova, Stanimira Yordanova

University of National and World Economy, Sofia, Bulgaria
dkabakchieva@unwe.bg, kstefanova@unwe.bg, smira.yordanova@gmail.com

Abstract. Business Intelligence is becoming very important for managers of companies with various sizes and operating in different industry sectors. Successful business leaders are well aware of the advantages that Business Intelligence technologies and tools provide for increasing their competitiveness and innovation in the globalizing and very dynamic business environment. The main purpose of this paper is to present the latest trends in the Business Intelligence system development. The main usage and technology trends are related to the users' increasing needs for easy and immediate access to data, at any place and at any time, which leads to the development of Mobile BI applications, Real-Time BI and Self-Service BI tools. Because of the constantly increasing volumes and radically different data, new Data Discovery and Exploration tools are developed and provided by vendors. In-Memory technology, delivering Business Intelligence and Analytics in the cloud, and Social BI are also important trends that will influence the BI system development in the next years.

Keywords. Business Intelligence, Mobile BI, Self-Service BI, Cloud BI, BI Trends

1. Introduction

Business Intelligence (BI) is becoming very important for managers of companies with various sizes and operating in different industry sectors. Successful business leaders are well aware of the advantages that Business Intelligence technologies and tools provide for increasing their competitiveness and innovation in the globalizing and very dynamic business environment.

The urgent business needs to make quickly strategic managerial decisions are well supported by Business Intelligence systems that are turning data into actionable information and providing visibility for effective corporate performance management. This is proven by the latest surveys in the IT field, conducted by world leading technology research and consulting companies like Gartner Inc., Forrester Research Inc., TDWI, etc., revealing that Business Intelligence and Analytics are among the fastest growing sectors in the overall worldwide enterprise software market.

The main purpose of this paper is to present the latest trends in the Business Intelligence system development. The main usage and technology trends are related to the users' increasing needs for easy and immediate access to data, at any place and at any time, which leads to the development of Mobile BI applications, Real-Time BI and Self-Service BI tools. Because of the constantly increasing volumes and radically different data, new Data Discovery and Exploration tools are developed and provided by vendors. In-Memory technology, delivering Business Intelligence and Analytics in the cloud, and Social BI are also important trends that will influence the BI system development in the next years.

2. Main Aspects of BI Systems Development

BI Systems are becoming more and more popular, and develop rapidly during the last decade. The BI software market experiences a lot of changes and confronts many challenges. Three periods are considered when discussing BI systems development aspects, based on the available research and analyses conducted by world leading technology research and consulting companies.

In the period between 2005 and 2008 [1], surveys among IT users from different industry sectors reveal that the implementation of Enterprise Resource Planning (ERP), Customer Relationship Management (CRM) and Supply Chain Management (SCM) solutions is not sufficient for achieving high efficiency and effectiveness. Companies are looking for systems ensuring optimal business processes and operations in order to stay competitive. The availability of large volumes of data requires suitable IT decisions and support and the BI systems attract the attention of a larger number of business managers. What is also peculiar in this period is the fact that BI system development projects are initiated not only in large enterprises, but also in small and medium-sized companies. The BI vendors also focus their efforts on the development of solutions suitable for SMEs, offering BI systems based on the SaaS model.

Business consolidation is another typical aspect of the BI software market, with the absorptions of Hyperion by Oracle, Cognos by IBM and Business Objects by SAP. The BI open source community also becomes more active and a number of pilot projects are started leading to increased competence pressure over traditional BI vendors. The great business potential of Data Mining technologies is realized and therefore many BI vendors start to include such functionalities in their BI system offers, either by developing Data Mining tools of their own or by acquiring smaller IT companied working in that field. New price models are also offered to users, e.g. providing licenses based on the number of users or the number of servers used.

During the period 2008-2011, the worldwide market for Business Intelligence is still rapidly growing in spite of the difficult economic situation. According to Gartner's annual global CIO survey in 2011 [2], BI software is ranked number five on the list of the top 10 technology priorities for chief information officers (CIOs). However, the business users are already demanding easier to use and more flexible BI solutions, while the IT experts need standardization and control. Gartner's experts talk about the "consumarization" of BI, referring to the needs of providing simple, intuitive and interactive BI tools, that could be used on mobile devices like smart phones and tablets, and "provide fun", in order to attract more business users. Another challenge for BI systems in this period is the need for handling large volumes of data from diverse sources, leading to the appearance of in-memory technology and social and content analytics.

In 2012-2013 the BI market continues to expand. Analytics and BI solutions are ranked number one technology priority for Chief Information Officers (CIOs) in Gartner's 2012 CIO survey [3]. The leading BI vendors continue to consolidate and dominate the BI market. However, a number of innovative vendors also appear and have a strong potential to establish themselves on the market. In response to the business users' needs, self-service data discovery tools and packed analytics to provide business context appear. Traditional on-premises solutions are still dominating the BI market but mobile and cloud technologies are becoming more important and get more attention. Delivering analytics in the cloud and mobility becoming essential are also among the five top BI trends [4] determined by the Data

Warehousing Institute (TDWI), a world leading organization providing education, training and research for executives and IT professionals in the Business Intelligence field.

3. Main Factors Influencing BI System Development

The most important factors that drive the BI system development include new technologies, rapid changes in the generated data, and more demanding information users. The lifecycle of new technologies is becoming shorter which makes them cheaper and consequently, more affordable for data collection, storage, access and analysis. The appearance of cloud computing also provides possibilities for a cheaper access to new BI technologies because organizations don't even have to buy new hardware and software but rent them at affordable prices. The development of mobile computing makes possible the easy and immediate access to data at any place and at any time.

According to leading experts in the BI field, the next generation of Data Warehousing and Business Analytics will mostly be influenced by "Big Data". The term "Big Data" refers not only to the enormous volumes of data that are generated nowadays, but also to the great variety of data, requiring extremely complex data management and analytics technologies, and to the increasing data velocity. The data volume is already measured not only by the quantity of transactions, events or history amount, but also by the number of dimensions, attributes, predictive variables. The data velocity refers to the continuously increasing speed at which data is created, accumulated and processed, which leads to demanding real-time processing and decisions. The data variety also changes quickly and therefore requires sophisticated methods and tools for storing and analysis. Over the past 20 years data has increasingly become unstructured, coming from sources beyond operational applications (producing predominantly structured data). Latest research shows that unstructured data tends to grow exponentially, while structured data tends to grow linearly, and moreover, unstructured data is vastly underutilized. The variety of data sources includes internet data (clickstreams, social media, social networking links), primary research data (experiments, observations, surveys), secondary research data (competitive and marketplace data, industry reports, consumer data, business data), location data (mobile device data, geospatial data), image data (video, satellite images, surveillance data), device data (coming from different sensors, RF devices, telemetry), etc., and continues to increase. All these changes in the available data lead to great complexities in data integration, transformation and processing.

The increasing demands of information users lead to convergence, i.e. to the need for merging traditional data management and analytics software and hardware technologies, open-source technology and commodity hardware, and creating new alternatives for IT and business executives to address Big Data analytics. New proprietary and open-source technologies enable different approaches for easier and more affordable storage, management and analyses of data. Hardware and storage is affordable and continuing to get cheaper to enable parallel processing. There is an increasing interest in handling the variety of unstructured data.

4. Current Trends in BI System Development

All of the above described factors lead to the development of Mobile BI applications, Real-Time BI and Self-Service BI tools. Because of the constantly increasing volumes and diversity of data, new Data Discovery and Exploration tools are developed and provided by vendors. In-Memory technology and Social BI are also important trends that will influence

the BI system development in the next years. Some of the most important trends in the BI system development are presented below.

Data Discovery and Self-Service BI

Information users nowadays are demanding an easy and immediate access to the data that they need for quickly taking managerial decisions. They are already not satisfied with traditional BI dashboards, providing preliminary defined analytical data, but prefer to be able to assemble their own dashboards and key performance indicators, and to get possibilities for linking insight to action. BI vendors are responding to these demands by providing the so called data discovery tools [5]. "Data discovery" is a new term used to describe the new wave of Business Intelligence that enables users to explore data, make discoveries, and uncover insights in a dynamic and intuitive way versus predefined queries and preconfigured drill-down dashboards. With data discovery tools the IT experts are providing the infrastructure, but the business people produce analytics and reporting, creating their own reports and dashboards. This is also referred to as Self-Service BI – an approach to business analytics that enables business users to access and analyze corporate data without the direct participation of IT experts who are involved mainly in the data warehouse and data marts set up and maintenance. The Self-Service BI approach is potentially beneficial for both groups, the end users being able to create personalized analytical queries and reports, and the IT staff being able to focus on tasks ensuring high quality data.

Mobile BI

Mobile BI starts with the appearance of smart phones in 1999. However, it is becoming mainstream with the launching of tablets, leading to the increasing demands of end users, insisting to have an easy and instant access to information in order to make managerial decision at any time [6]. The massive changes in the way of consuming information have also affected the development of BI applications. Mobile BI applications are easy to deploy, they often appear quite simple to the users and yet satisfying their business needs, and ensure 24/7 access to BI and data analytics. Nowadays there are two main strategies for delivering Mobile BI - purpose-built and web-based. The purpose-build Mobile BI applications are device-specific, e.g. iPhone or Android applications downloaded from iTunes or the Android Market. The web-based Mobile BI applications are using a web browser and work on most devices without an application being installed on the device. Mobile BI applications can increase the productivity of business users but there are some challenges to consider [7]. The main problems are related to data security concerns, the small screen size and limited memory of mobile devices.

Cloud BI

As already mentioned, the appearance of cloud computing provides possibilities for a cheaper access to new BI technologies because organizations don't even have to buy new hardware and software but rent them at affordable prices. Companies do not have unlimited capital to invest in infrastructure while their data is exploding and therefore the idea for using BI in the cloud appears. However, for various technical, political, social, regulatory and cultural reasons Cloud computing has not yet become a successful business model widely

adopted for enterprises to store their large data assets. The main obstacles to overcome before moving BI on the Cloud are related to security and trust [8]. Companies always keep their important data on premise, and the public cloud remains untrusted for many organizations.

Other Technological Trends for BI Development and Delivery

The number of BI vendors, delivering *open source* BI solutions, increases on the BI market. The open source BI software is becoming popular and well accepted, especially among small and medium-sized enterprises (SMEs) that cannot afford large budgets for expensive proprietary solutions. Open source BI is not constrained from proprietary vendors' ideas and vision, but allows users to develop applications in the way that suits best their business needs. Open source technology provides also flexibility in the adoption model – the BI software could be downloaded and deployed at any moment when users need it, the adoption could be carried out at the users' own pace, it could be further extended with new functionalities, and last but not least, it could be acquired at a lower cost.

In-memory BI is another technological approach used recently in BI systems to solve complex and time-sensitive business scenarios. Traditional BI deployments are usually disk-based, i.e. the application queries data stored on physical disks. With in-memory BI, the queries and data reside in the server's random access memory (RAM). The in-memory approach improves the overall speed of a BI system and provides business users with faster answers, especially for queries that take a long time to process in a large database.

Traditional BI solutions are very complex systems, considering multiple components including data warehousing and the associated problems of data integration and cleaning, query and reporting tools, OLAP, data mining analytic tools, metrics creation and management, dashboards, alerts, etc. [9]. For these reasons, traditional BI system development approaches are very complex and consist of many stages, steps and activities. The term *Agile BI* is recently used to refer to the agile software development methodology for BI project implementation. This different approach reduces the time-to-value of traditional BI and makes BI applications more flexible and able to react much faster to ever-changing business and regulatory requirements.

A new BI delivery model is also the *Software-as-a-Service BI* (SaaS BI), sometimes referred also as *on-demand BI* or *cloud BI*, in which BI applications are not implemented on the premises of a company but are employed at a hosted location and is accessed via protected Internet access. As a difference to the conventional software licensing model with annual maintenance or license fees, the SaaS BI generally implies a pay-as-you-go or subscription model. The SaaS BI delivery model main advantages are the easier access to solutions without building and maintaining an own onsite implementation, the immediate access to BI and predictive analytics, and it is a good choice when there is low budget for BI software and related hardware. The disadvantages, however, are the limited features if compared to on-premises software products.

Another term, associated with delivering BI solutions in the cloud, is *social BI*, which is either referred to as the creation and sharing of BI analytics reports and dashboards by end users of cloud computing, or as providing BI based on social networks data. Social BI is enabled by the rapid growth of social media networks in 2009 and will continue to play an important role in the BI field.

5. Conclusion

The BI market is growing and will continue to be on the rise. The Analytics and Business Intelligence solutions will still be a primary technology priority for CIOs worldwide, in all industry sectors and for companies of all sizes. The next generation BI will mostly be influenced by Big Data, big in volumes, variety and velocity, and the increasing demands of business users requiring an easy and immediate access to the data that they need for quickly taking managerial decisions, assembling their own dashboards and key performance indicators, and getting possibilities for linking insight to action. In response to these changes in the business environment, BI vendors will focus their efforts on providing Data Discovery and Self-Service BI tools, and Mobile BI applications. The market share of open source BI solutions will increase. In-Memory technology, the agile approach to BI system development, delivering Business Intelligence and Analytics in the Cloud, and Social BI are also important trends that will influence the BI system development in the next years.

References

1. CIO Bulgaria Journal (2008). 11 Tendencies in the BI platform development" (in Bulgarian). CIO Bulgaria Journal, Vol.5, 2008. Available at: http://cio.bg/1912_11_tendencii_v_razvitieto_na_platformite_za_biznes_analizi
2. Gartner Inc.(2011). Gartner Forecasts Global Business Intelligence Market to Grow 9.7 Percent in 2011. Press Release, Sydney, Australia, February 18, 2011 Available at: http://www.gartner.com/newsroom/id/1553215
3. Gartner Inc. (2012). Gartner Executive Programs' Worldwide Survey of More Than 2,300 CIOs Shows Flat IT Budgets in 2012, but IT Organizations Must Deliver on Multiple Priorities. Press Release, 18 Jan 2012. Available at: http://www.gartner.com/newsroom/id/1897514
4. TDWI (2012). 5 Macro Trends Shaping Next-generation BI and Analytics. Available at: http://tdwi.org/articles/2012/02/07/5-macro-trends-bi-analytics.aspx
5. CGI Group Inc. (2011). Next Generation Business Intelligence. Available at: http://www.cgi.com/sites/default/files/white-papers/business-intelligence-white-paper.pdf
6. Search Business Analytics. Mobile BI tools, trends and best practices guide. Available at: http://searchbusinessanalytics.techtarget.com/essentialguide/Mobile-BI-tools-trends-and-best-practices-guide
7. Parizo, C. (2013). Growing use of mobile BI apps sparks need to downsize designs. Available at: http://searchbusinessanalytics.techtarget.com/feature/Growing-use-of-mobile-BI-apps-sparks-need-to-downsize-designs
8. Tata Consultancy Services Ltd. (2012). White paper, Business Intelligence on the Cloud: Overview and Use Cases. Available at: http://www.tcs.com/SiteCollectionDocuments/White%20Papers/HighTech_Whitepaper_Business_Intelligence_Cloud_0412-1.pdf
9. Evelson, B. (2010). Agile BI Out Of The Box. Available at: https://www.wherescape.com/getattachment/resource-library/white-papers/agile-bi-out-of-the-box/Agile-BI-out-of-the-box.pdf

Enhancing Business Intelligence with Sentiment Analysis

Stanimira Yordanova[1], Kamelia Stefanova[2], Dorina Kabakchieva[3]

[1] UNWE, 1700 Sofia, Bulgaria
smira.yordanova@gmail.com

[2] UNWE, 1700 Sofia, Bulgaria
kstefanova@unwe.bg

[3]UNWE, 1700 Sofia, Bulgaria
dkabakchieva@unwe.bg

Abstract. The present paper presents Business intelligence (BI) concept of analysing data coming from structured and unstructured sources to improve decision making process in organization. The sentiment analysis for extracting opinion from unstructured content is introduced to convert unstructured customers` opinions into meaningful structured data, so the data can be further analyzed in BI applications in order to identify trends and patterns.

Keywords. Sentiment Analysis, Opinion Mining, Business Intelligence

1. Introduction

Nowadays companies face the challenge of explosive growth of both structured and unstructured data. While 20% of the data available in companies is structured, the other 80 % is unstructured. Unstructured data are all text documents, logs, survey results, and e-mails within the company; external data like social media data, generated in Facebook, Twitter, LinkedIn, and Flickr; mobile data including text messages and location information and website content like forums, blogs, review websites and etc.

Companies are challenged to collect, understand and use internal and external data to make proper and timely management decisions in order to meet the constantly changing market requirements. Trying to respond to the strong competition, they need to be continuously focused on increasing customer satisfaction by improvement and extension of their product offerings. To succeed, companies need deeper insight into the current customer experience. They usually organise special surveys for identification of the customer opinion on the products, offer them various incentives, collect and analyze received results. Such method of gathering customers experience is proven to be time-consuming and expensive.

Usage of social media networks and collaboration technologies becomes essential part of people lives by providing an opportunity to post their opinion about a company`s product and services in a form of reviews, comments, ratings. This leads to storing a large quantity of vital information relating to a customer in the form of unstructured text containing customer`s opinion in the review websites. To gain insight into the customer experience and opinion, it is important for companies to aggregate and combine structured and unstructured data to perform correlations and advanced analytics.

Opinions expressed and posted on the review websites by customers are natural language text. Text analysis techniques are used to extract meaningful information and

transform it into structured data, thus deriving potentially valuable business insights from unstructured textual content. Sentiment analysis uses text analytics in order to extract the opinion on the author from a review, converts unstructured subjective content of textual customers` opinions into structured data so the data can be analyzed in Business intelligence (BI) applications in order to identify trends and patterns and to improve the decision making.

Applying sentiment analysis methods in Business intelligence systems, described as main emerging Business intelligence trend [8] allows companies to know the opinion of their customers and better target their campaigns by determining the positive or negative sentiment for their products or services in the products` reviews.

2. Business Intelligence

According to the definition of Gartner [1] "Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance".

The TDWI [2, 8] defines BI as "The processes, technologies and tools needed to turn data into information, information into knowledge, and knowledge into plans that drive profitable business action. Business intelligence encompasses data warehousing, business analytic tools, and content/knowledge management."

Deborah Quarles van Ufford [3] presents BI as a pyramid with three layers. Data Warehouse is the basis and on top of it are Queries and Reports, On-Line Analytical Processing and Data Mining.

There are many other definitions for Business intelligence which are given by BI experts, researchers and vendors. The main focus is on:

- Business data is coming from diverse type of data sources, including structured and unstructured content;
- Data Warehouse is the significant component of the BI, as it stores transaction and non-transaction data transformed to be suitable for querying, reporting and analysis;
- Business intelligence is not just a technology. It encompasses people and knowledge as well. The real value of the BI is in delivering knowledge to the right people willing to make decision and to take profitable action.

Business intelligence systems use data from transactional systems aggregated and stored in Data Warehouse. Extraction, transformation and load (ETL) processes are responsible for incorporating the enterprise data in a Data Warehouse. Analytical Models (On-line Analytical Processing (OLAP), Data Mining and Analytical Reporting) are used for processing the data. Sophisticated graphical instruments for identification of complex relationships, reporting and presentation of data provide users with information in a comfortable and accessible form (Fig.1)

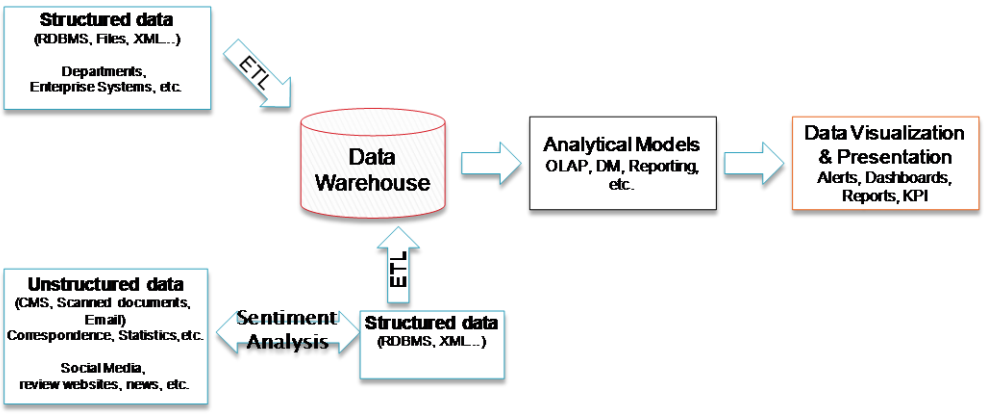


Figure 1: Business Intelligence Model

The quality of BI analytics can be improved by including unstructured data in the analysis resulting in new types of insights for the BI users. Unstructured data cannot be directly analysed by Business intelligence tools. A BI system can process the unstructured data after analysis and reducing of the unstructured data and producing a data model using specialized software for text analysis. Text analysis tools extract information such as product names, product features and customer sentiment from the unstructured data and introduced it to the ETL processes that feed the data warehouse. The extracted information can be integrated into the existing warehouse schemas and cube definitions. Analytical models are used to explore the data and to present it in a useful and proper manner to the end user.

Sentiment analysis uses natural language processing, text analysis and computational linguistics to identify and extract subjective information [9] (i.e. product features and opinion polarity) in textual content, and converting unstructured subjective content of textual customers` opinions in reviews into structured data for further analysis in a BI system.

Enhancing Business intelligence with Sentiment analysis can help organizations to become more active in addressing negative reactions to products before they lead to poor sales that BI users detect later in the reporting and analysis of sales transaction figures.

3. Sentiment Analysis Definition

Sentiment analysis is also called opinion mining became popular research field after 2000, because of web 2.0 technologies explosion when people start to share their opinions on the web. Nowadays, it is researched in data mining, web mining, and information retrieval.

By the definition given by B. Liu and L. Zhang [4], "sentiment analysis is the study of people`s opinion, appraisals, attitudes and emotions toward individuals, organizations, issues, events topics, products and their attributes". Sentiment analysis uses techniques from text analytics to identify and extract opinions from unstructured human-authored documents [5]. "Opinionated text" [6] in the websites is the object of the sentiment analysis aiming at extraction of the opinion expressed by the author in the text called sentiment classification.

Sentiment classification is performed at three levels as described in the [4,6]: document-level, sentence level and aspect level classification.

Document level classification classifies the whole document as positive or negative while the sentence level determines the sentiment polarity of each sentence and defines it as objective or subjective. In order to determine the polarity of the sentiment about a particular aspect of a product, aspect level classification is needed.

4. Aspect based Sentiment Analysis

Aspect based sentiment analysis includes extraction of aspects from a review and aspect level classification of the opinion on the extracted aspect. As a result of this analysis unstructured text is turned into structured data comprised of product, its features, author`s opinion about them classified as positive or negative.

The opinion components in a product review usually are the author, the date of the review, the product feature and the text with sentiments about a product/service. Each product is presented in the review as a hierarchy of components and sub-components [4, 7]. The target of an opinion can be the product or its features (also called in the research literature "aspects". The author can express his/her sentiment about the product or/and its feature. Opinion components identified and extracted can serve as a base to determine a database scheme.

After crawling and storing, the text corpus is analysed including the following steps: aspect extraction; sentiment classification and opinion summarization (Fig.2). All opinion components are identified by performing extraction of aspects (words and phrases that appear in the text and are variations of the aspect category), and grouping synonymous aspect into aspect category. Aspect sentiment classification determines the opinion polarity of the extracted aspect.

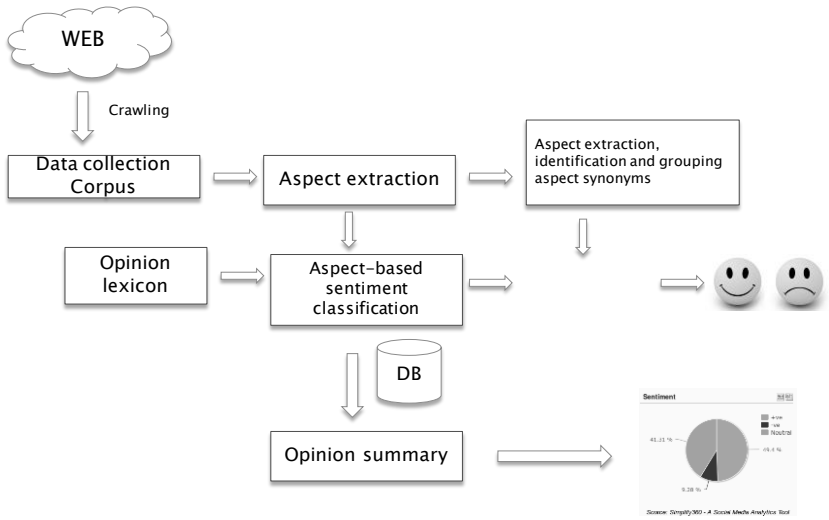


Figure 2: Aspect based sentiment analysis process

There are many different approaches and methods described in the research literature about aspect extraction and aspect sentiment classification because it is a very active research field in the last years. Aspect extraction methods are based in different approaches – using supervised learning methods (i.e manual sequential labeling; Hidden Markov Models,

Conditional Random Fields and etc), unsupervised learning methods (i.e topic modeling) and data mining algorithms [12] for extracting relations between opinion and its target.

Hu and Liu (2004) uses POS tagging for finding explicit aspect expressions and frequency counting for extracting frequent nouns and noun phrases from reviews in a domain. If a sentence contains sentiment word, the nearest noun or noun phrase to the sentiment word is extracted. This approach is based on the finding that people uses different vocabulary in commenting aspect and relies on identification of aspects because people commented them frequently. It is accepted as an effective method and many other researchers work on the frequency based approach to extract aspects. As described by Bing Liu, the method was improved by removing noun phrases that are not aspects using pointwise mutual information (PMI) score [13]; applying filters (Blair-Goldensohn, 2008); using frequency-based approach with pattern-based filter (Moghaddam and Ester, 2010). [6]

Clustering analysis methods and similarity metrics based methods to combine aspects discovered into groups in a taxonomy are used to identify and group synonyms of aspects.

Supervised [10] and unsupervised [11] learning methods are applied for aspect sentiment classification.

Turney`s algorithm uses POS tagging for labeling each word in a sentence with its appropriate part of speech and extracts phrases containing adjectives or adverbs as opinion indicators. Measuring statistical dependence between two terms by PMI measure, the sentiment polarity of extracted phrases is determined:

$$PMI(term_1, term_2) = \log_2\left(\frac{\Pr(term_1 \wedge term_2)}{\Pr(term_1)\Pr(term_2)}\right) \quad (1)$$

$\Pr(term_1, term_2)$ is the actual co-occurrence probability of $term_1$ and $term_2$,
$\Pr(term_1)\Pr(term_2)$ is the co-occurrence probability of the two terms if they are statistically independent. The ratio between both is a measure of the degree of statistical dependence between them.
The sentiment orientation of a phrase is computed based on its association with the positive reference word “excellent” and the negative reference word “poor”:

$$SO(phrase) = PMI(phrase, \text{“excellent”}) - PMI(phrase, \text{“poor”}) \quad (2)$$

Ding, Liu and Yu (2008) introduced lexicon-based method, comprised of four steps: marking sentiment words and phrases, apply sentiment shifters, handling but-clauses and aggregating opinions. The lexicon contains sets of sentiment words, which may have different orientations in different domains, which makes lexicons domain dependent.

Supervised learning methods are dependent on the training data in one domain, which makes them difficult to adapt in another domain. While applying unsupervised learning methods using lexicon-based approach can be applied for different domains if the lexicon is complied. Such sentiment lexicons are compiled manually by researchers using set of seed positive and negative words prepared separately for adjectives, adverbs, verbs and nouns. Using WordNet - a lexical database for English (http://wordnet.princeton.edu/) each seed is enriched with synonyms and antonyms and the incorrect words are manually removed. There

are some public domain sentiment lexicons such as SentiWordNet[1], Bing Liu's sentiment lexicon[2], Emotion lexicon[3].

5. Conclusions

New collaboration technologies allow customers to express dissatisfaction or satisfaction from bought product very quickly on review websites or social networks. Business is challenged to respond quickly to the information coming from their customers as they struggle to build consumers loyalty. By extracting opinion of the customers on the web, sentiment analysis can provide meaningful business intelligence identifying problems with products and services.

Many researchers are experimenting with different methods for sentiment analysis. Aspect based sentiment analysis suggests techniques for extraction and classifications resulting in structure generation from the unstructured web content. This provides many opportunities to analyze the structured data obtained through sentiment analysis in Business intelligence applications. Integrating data extracted from the customers` reviews with business data and applying Business Analytics tools to identify new relationships between the data or make predictions is new emerging trend in the research field and in vendors' products.

References

1. Gartner, Inc., http://www.gartner.com/it-glossary/business-intelligence-bi/
2. The Rise of Analytic Applications: Build or Buy? by Wayne W. Eckerson, TDWI reports (2002)
3. Quarles van Ufford, Deborah. Business Intelligence: The Umbrella Term, VU Amsterdam (2002)
4. Bing Liu, Lei Zhang, A survey of opinion mining and sentiment analysis, Mining Text Data (2012)
5. Bo Pang and Lillian Lee, Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval, Vol. 2, Nos. 1–2 (2008) 1–135, 2008
6. Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012, http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
7. Bing Liu, Web Data Mining, Exploring Hyperlinks, Contents, and Usage Data, Springer Berlin Heidelberg New York, 2007
8. David Loshin, Business intelligence: The savvy manager's guide – 2nd ed. (2012), http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
9. http://en.wikipedia.org/wiki/Sentiment_analysis
10. Peter D. Turney, Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002, pp. 417-424.http://acl.ldc.upenn.edu/P/P02/P02-1053.pdf
11. X. Ding, B. Liu, P. Yu, A Holistic Lexicon-Based Approach to Opinion Mining, Proceeding WSDM '08 Proceedings of the 2008 International Conference on Web Search and Data Mining, pp. 231-240, http://www.cs.uic.edu/~liub/FBS/opinion-mining-final-WSDM.pdf
12. M. Hu, B. Liu, Mining and summarizing customer reviews, Proceeding, KDD '04 Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 168-177, http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
13. Popescu and O.Etzioni, Extracting Product Features and Opinions from Reviews, http://turing.cs.washington.edu/papers/emnlp05_opine.pdf

[1] http://sentiwordnet.isti.cnr.it/
[2] http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
[3] http://www.umiacs.umd.edu/~saif/WebPages/ResearchInterests.html#SemanticOrientation

Advantages of Using a System for Intelligent Tagging and Search in Unstructured Data

Maria Marzovanova[1]

[1] UNWE, 1700 Sofia, Bulgaria
mmarzovanova@unwe.bg

Abstract. The system for intelligent tagging and search based on extended taxonomic structure aims to assist the processes of web information storage and retrieval. The number and the features of such type of software solutions is far from being a small one. However, the system presented in this paper has two differentiating features. First, it offers a new technology for information organization using a newly developed structure – extended taxonomic structure, and second, it provides means for accessing various text documents' formats in order to interpret their content and receive reliable search results.

Keywords. Tagging, search, extended taxonomic structure, text documents

1. Introduction

Knowledge organization systems are intended to help search engines by applying diverse content arranging techniques. Search engines and their job have been object of many studies for quite a long time and therefore numerous methods and solutions for improvement are available.

There are two main kinds of knowledge organization systems depending on the base structure – based on taxonomies and based on folksonomies. Some of their main features are briefly presented in table 1.

Table 1

Taxonomy	Folksonomy
• usually part of a specific system	• user-generated
• created by specialist	• no control and predefined structure
• difficult to customize and with strong rules for modification	• Categorizing web data

Analyzing their application principles, it's easy to notice the compromises that a system does with its results in order to gain in some advantages. The logical approach in search technology's development is to try to combine these two technologies' advantages into a new one. There are some works in this area too.

Ferdy Christant, for example, has developed a system called Image Dragon [1]. He applies some taxonomy based features onto a free-tagging system for photos. In this way he achieves the possibility to imply hierarchy in the categories and extract additional context information from that.

Ching-Chieh Kiu and Eric Tsui present another similar result – algorithm for creating a hybrid taxonomy-folksonomy [2]. They describe four phases of transforming a folksonomy to taxonomy.

The structure lying in the foundation of the system for intelligent tagging and search is extended and customized taxonomy [3]. And the focus here is not on changing or refining the basic structure, but making the information in it wider and deeper. The basic layer consists of taxonomy of terms extracted from a trusted source, which gives some degree of control and reliability. This skeleton taxonomy is afterwards enlarged by adding new terms and diverse relationships between the terms. The richer vocabulary base is the main feature that allows the implementation of different mechanisms for indexing of the available sources of information and refining the search results from them.

When talking about information retrieval systems important is not only the building of the structure of terms to be used for managing the available information. Another aspect of the problem is using the resultant structure for easier and more effective work of the system.

Both approaches mentioned above for categorizing information, taxonomy and folksonomy, demand user intervention. Putting a piece of information in the right category as well as retrieving only the relevant information after accomplishing a search action need the user to read the source content. The system for intelligent tagging and search implements two methods, which automate the processes of categorizing content and evaluating results based on the structure of terms. [4]

2. Purpose

The purpose of the report is to demonstrate the main features of the system for business intelligent tagging and search. The focus is on the processes of intelligent tagging using terms from a taxonomic structure and business intelligent indexing of the found sources according to the logic implemented while building the taxonomic structure.

The preparation for the final result consists of building two taxonomic structures in different thematic fields – economics and informatics, adding some text documents to the system's repository and, on that foundation, performing tagging and indexing procedures. Then a search of specific term included in both of the taxonomic structures will prove the difference in the system's interpretation of the term depending on the chosen taxonomy.

If we take the term stock, for example, and look up that word in a dictionary there are a number of meanings listed and illustrated. Some appropriate for the experiment definitions of the word, given in Longman Dictionary of Contemporary English [5], are:

FINANCE

a) [countable]*especially American English*aSHARE in a company:
the trading of stocks and shares

b) [uncountable]the total value of all of a company'sSHARES

6

COOKING

[uncountable and countable] a liquid made by boiling meat or bones and vegetables, which is used to make soups or to add FLAVOUR to other dishes:

chicken stock

vegetable stock

3. Taxonomies

For the experiment we will use parts of two different taxonomies. In order to compare the search results we have to choose a term that appears in both of the taxonomies. Having this in mind, appropriate for the test procedure is the term "stock" with its meaning in the area of cooking and its other meaning in the finance's vocabulary. As the term is exactly the same in both cases, the purpose of this test is to show how depending on the terms connected with the chosen one, when searching using the one or the other, the term is interpreted differently. The logic implemented in the application is that if a specified number of the related terms, three in this case, are found in the same text document as the searched one, we can assume that the text is among the wanted results.

In the first classification structure, taxonomy of cooking terms, stock is related to terms like pan, vegetables, meat, herbs and water (Figure 2). All these terms are meant to explain the meaning of stock in the current taxonomy position aspect. This means that if we find words meaning pan, vegetables, meat, herbs and water, the text is very likely about stock in the aspect of flavoured water and we are in the field of cooking.

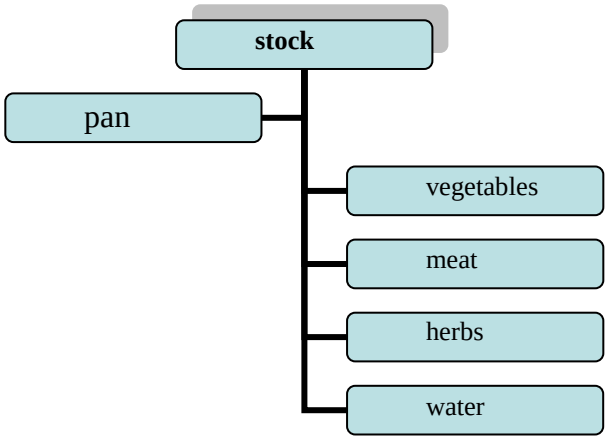


Figure 2

In the second taxonomy structure the term is located on the second level. Its parent term is stock and the terms related to it are markets, risk and capital, insurance and shares (Figure 18). This means that if we find in a text document anything related to three of these five terms the text is about „stock" in the aspect of finance.

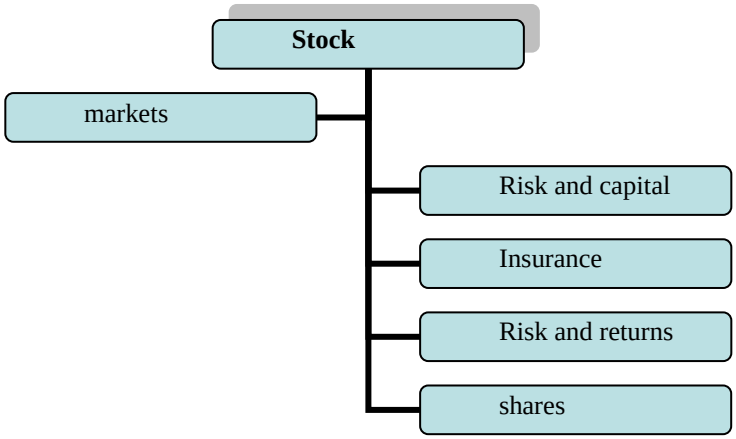


Figure 3

4. Text documents

For testing purposes there are three files added to the system's repository. In order to demonstrate the system's ability to work with different file formats, they are chosen to be txt, docx and pdf. In terms of contents there is one file related to cooking, one file related to finances and one file containing the word stock with different meaning.

The first document is titled Vegetable stock, it's in docx format and contains three of the taxonomy terms (Figure 4).

Vegetable stock

Ingredients

- 2 carrots, peeled and roughly chopped
- ....
- 3 fresh parsley stalks

Preparation method

1. Put all the vegetables and the other ingredients into a large pan and cover with water. Bring to the boil. Cover and simmer very gently for 20-30 minutes. Strain into a large bowl and allow to cool.
2. Use within 3 days or freeze....

Figure 4

The second document is titled Stock market and is in txt format. It contains four of the terms used in the taxonomy branch of the term „risk and capital" from the finance's structure (Figure 19).

Stock market

A stock market or **equity market** is the aggregation of buyers and sellers (a loose network of economic transactions, not a physical facility or discrete entity) of stocks (shares); these are securities listed on a stock exchange as well as those only traded privately.....

Market participants include individual retail investors, institutional investors such as mutual funds, banks, insurance companies and hedge funds, and also publicly traded corporations trading in their own shares....

Figure 5

And the third document which is in pdf format is titled Product settings. One usages of the only term among the mentioned in the taxonomies, stock, is highlighted in Figure 6

Stock

Stock Status	Manually increment (by entering a positive quantity) or decrement (by entering a negative quantity) the stock quantity of a product.The quantity you enter (which can also be zero), will be automatically updated as this product is shipped or returned.....
Do Not Allow Back Orders	For products that can be back-ordered, stock status can reach negative values if the product is ordered after its stock value has reached zero...
Link Stock Status with ProductCode	Enter another product's code to have this product's stock status shared with that product to link product stock from one product to another. This is often used to offer a free accessory with purchase of one product so that the stock of one product never exceeds the stock of the free accessory.

Figure 6

5. Tagging and indexing

There are two groups of actions according to the implemented in the system methodology that have to be performed before the actual search [4]:

- tagging the sources;
- indexing the tags.

These two methods are fully automated and require no user interaction. The automated execution is based on the logic lying in the taxonomy. The tagging process includes as a first step preparing a list of all the terms participating in the taxonomy. Each element is then annotated in the document as appears and added to the list of tags in the system. With this the first step is finalized.

The indexing process consists of matching the added annotations back with the taxonomy elements and creating pointing indexes, connection between the annotations in text with the elements in the classification. As mentioned above for testing purposes there is a set rule that a term is considered relevant only if there are at least three of its connected terms in the same text. Then, an index in the system is added only for those terms.

With tagging and indexing performed all the preparation is finished. All the added sources are annotated and all meaningful uses of the terms are indexed in the system.

All taxonomy terms used in the documents are annotated and as a result there is a list of annotations in the system. The resultant table contains two important columns for the current discussion – AnnotationTerm and DocumentUrl. Even though the table doesn't contain all the records, there are the URL addresses of the three files added to the system's repository with some terms found there. This illustrates two main points:

- all the files are accessible from the system despite of their different format,
- and all the files contain some of the terms participating in the taxonomies.

6. Results

The system provides a single-term search interface, which is very appropriate for the current experiment. The results of searching the term „stock" contain all the constructions having that word as a term or as a part of the term. Among all of them currently important are the two constructions discussed above (Figure 7,Figure 8). You can also see the searched term in green.

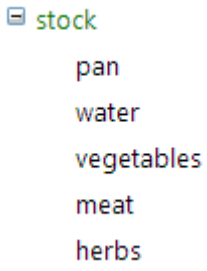


Figure 7

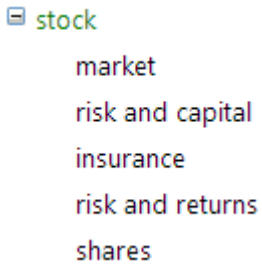
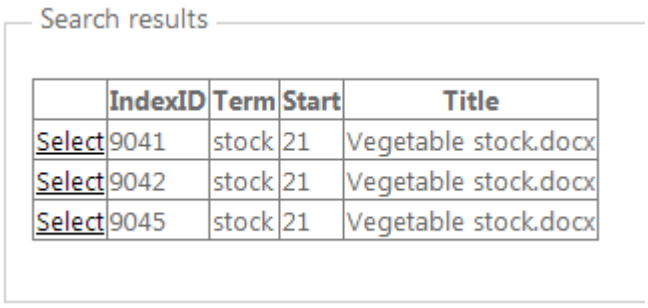


Figure 8

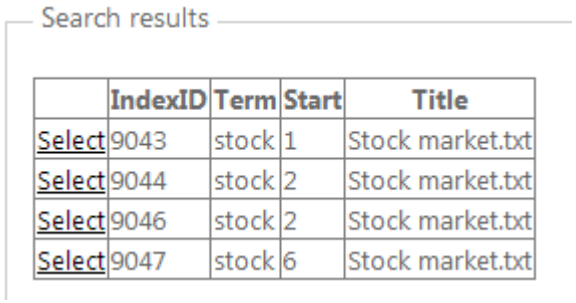
If the first usage is selected, this means that the user is working in the fields of economics and more specifically the wanted results are those concerning the meaning of the term - demand. The expected results, as indexed related documents, contain the title – Vegetable stock.docx.


Search results

	IndexID	Term	Start	Title
Select	9041	stock	21	Vegetable stock.docx
Select	9042	stock	21	Vegetable stock.docx
Select	9045	stock	21	Vegetable stock.docx

Figure 9

The second usage leads to a result containing indexes of the document – Stock market.txt


Search results

	IndexID	Term	Start	Title
Select	9043	stock	1	Stock market.txt
Select	9044	stock	2	Stock market.txt
Select	9046	stock	2	Stock market.txt
Select	9047	stock	6	Stock market.txt

Figure 10

Both tables with results provide access to the related documents. Selecting an index opens a page with the text of the document where the searched term is coloured in red for the user's comfort.

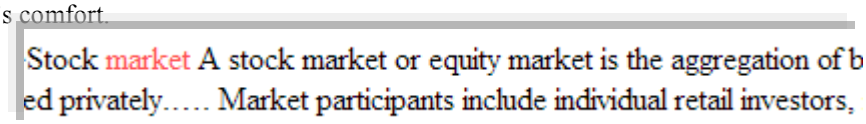
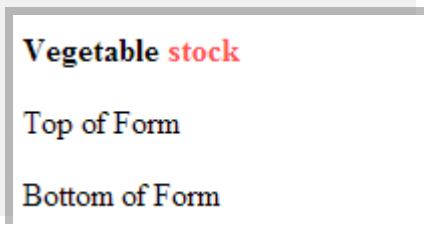

Stock market A stock market or equity market is the aggregation of b
ed privately..... Market participants include individual retail investors,

Figure 11


Vegetable stock

Top of Form

Bottom of Form

Figure 12

If we examine the indexed documents as a result of this experiment there are two main points to be made:

- only two of the three added document are indexed,
- each of the indexed documents is indexed only in relation with one of the taxonomies, although they both contain the searched term.

The missing document in the indexes is explained with the fact that it contains only the searched term and none of the related to it in any of the taxonomies. According to the set rule that a document must contain at least three of the related terms, this document becomes irrelevant. This rule explains the different documents indexed for the different taxonomies, too. Considering the system's methodology of indexing it can be concluded that each of these constructions has the searched term in different context meaning.

7. Conclusion

The presented example of the usage of intelligent system for tagging and search illustrates some of the main advantages of this approach. Automating the tagging process shortens the user's efforts and integrating the taxonomy logic in the search process provides some intelligent interpreting done by the machine, which makes the search results more accurate, as well as fast and easy to receive.

References

1. F. Christant, „Tagonomy - Taxonomy meets Folksonomy," 2010. [Онлайн]. Available: http://www.ferdychristant.com/blog//articles/DOMM-86CCFU.
2. C.-C. K. a. E. Tsui, „TaxoFolk: A hybrid taxonomy - folksonomy classification for enhanced knowledge navigation," Knowledge Management Research & Practice, том 8, pp. 24-32, 2010.
3. M. Marzovanova, „Agile Taxonomy/Folksonomy of Economics in Bulgarian Language," в UNITE 2nd Doctoral Symposium:R&D in Future Internet and Enterprise Interoperability, Sofia, 2012.
4. A. M. Marzovanova, T. Tsoykova и P. D. V. Kisimov, „INFORMATION SYSTEM FOR TAGGING, INDEXING AND SEARCH OF UNSTRUCTURED DATA," в Международна конференция „Приложение на информационните и комуникационните технологии и статистиката в икономиката и образованието – 2012", 5-6 Октомври, Sofia, 2012.
5. Longman Dictionary of Contemporary English, http://www.ldoceonline.com/

Formal Methods for Conflict Detection During Multi-Dimensional Data Mart Integration

Geno Stefanov[1]

[1] UNWE, 1700 Sofia, Bulgaria
genostefanov@unwe.bg

Abstract. A company Data Warehouse is usually built following a bottom-up approach, where the Data Warehouse is a set of materialized views over the whole information system of the organization, namely the Data Mart. The problem of integrating autonomously developed data marts can arise when an organization needs to merge data resides in independently developed data warehouses in order to make a better decision in a given situation. During this integration several categories of conflicts can arise, because of the complexity in the multi-dimensional data model on which a Data Mart is based. The multi-dimensional model introduces, in addition to the ER data model, dimension and fact entity. As a result of the multi-dimensional model elements there are two levels of heterogeneities – dimension and fact. The former depends on differences between both dimensions' hierarchies, members of the dimensions, and names of members, levels and dimensions. The latter kind of conflict occurs when facts in different Data marts are in different names, values (inconsistent measures), and formats or even on a different scale. This paper proposes formal algorithms (methods) for the detection of conflicts and heterogeneities which can arise when independently developed data marts are being integrated. The first section of the paper presents the literature overview on the topic. Then the proposed novel methods are defined and explained.

Keywords. Data warehouse, data marts, integration conflicts, data integration.

1. Introduction

Data Warehouses (DWs) are analytical information systems, i.e. databases that are optimized for analytical decision support rather than for high transactional throughput. More specifically, "a Data Warehouse is a subject oriented, integrated, non-volatile and time variant collection of data in support of management's decisions" [1]. Thus, a DW represents an integrated collection of historical business data, supporting business analysts in the decision making process. In that sense, the DW represents the "corporate memory" of all business data that is relevant for strategic decisions [2].

One problem that needs to be solved in many practical cases is the integration of data marts that have been developed and operated independently. In many cases, data which resides in multiple and independently developed data sources are needed for decision-making. For example, if we want to compare several KPI from different DM or to define new KPI, which is defined by the KPIs residing in several different DMs, we face two

possible choices – building a new DM or the integration of the existing ones. Building a new DM is a costly and time consuming task, therefore it would be better if we have mechanisms for integration of the existing DMs. Another case where DM integration is needed is in mergers and acquisitions of different companies. In this case for example one company acquires another company and the DMs in the acquired company should be integrated into the DW of the acquiring company. Here we have the same possible choices as in the previous case.

The multidimensional model (MDM) on which the concept of DW and DM are based is built up of three basic constructive elements: the facts which are analyzed, the dimensions (coordinates of the fact) and the measures which allow for the quantitative evaluation of the facts [3]. Very often in large organizations the need for integrating independently developed DMs arises. These data marts should be based on common dimensions and facts, but in many cases different departments of one company develop their own data marts and their integration becomes a difficult task. The difficulties come out of the heterogeneities of the MDM elements and can be classified as dimension conflicts and fact conflicts from a semantic point of view. The dimension conflicts occur when the dimension schema structures, dimension members or the naming of semantically related dimensions have semantic discrepancies. The fact conflicts occur when the measures in different DMs are in different names, values (inconsistent measures), formats or even different scale. Some work has been done on the problem of resolving the conflicts occurring in data mart integration [4], [5], [6] and [10]. The main part of the DMs integration is resolving the possible conflicts. To the best of our knowledge there are no proposals for methods or techniques which will allow the DMs integration process to automatically identify the possible heterogeneities between the integrated MDM models.

The goal of this paper is to define formal methods for conflict detection during the DMs integration process.

The rest of the paper is organized as follows: Section 2 introduces the analysis and summarization of the DMs schema elements and semantic conflicts occurring during the DM integration. In Section 3 two algorithms for conflict detection are presented and explained. Finally in section 4, some conclusions and future work are inferred.

2. Data Mart Construct and Heterogeneities

Intuitively, a DM defines one or more measure variables within fact tables, categorized by some dimensions that are organized in hierarchies of levels. Facts and dimensions consist of both their schema and the corresponding instances. For the dimension instances, we use the commonly accepted term dimension members (or members for short) [9] throughout the paper

In order to achieve the goal of the current paper, we have to analyze and define the constructing elements of one Data Mart. A DM = $\{F_1,..., F_n, D_1, ..., D_m\}$, consists of non-empty set of *dimensions* and *fact tables* [6], [7].

Now let $\{\tau 1, ..., \tau m\}$ be a finite set of data types (e.g.integers) with their domain defined by function dom(τ).

A dimension $D \in \{D_1, ..., D_m\}$ of DM is composed of:

- The dimension schema $S_D = \{L_D, S(L_D), H_D\}$ containing finite, non-empty set of Levels $L_D = \{l1,\ldots., l_m, l_{All}\}$, with level schema $S(L_D) = \{Sl_1,....,Sl_m\}$ and the roll-up hierarchy $H_D \subseteq L_D \times L_D$.

- The level Schema Sl ∈ S(L_D) of some level l_i is an attribute schema (K,N_1, ..., N_k) with name l, key K(the dimensional attribute) and optional non-dimensional attributes $N_{1,...}$ N_k. Every N_k attribute ∈ Sl has a domain dom(N_k) = dom(τ_k).
- The dimension instance d(S_D) over schema S_D with name d containing a set of members Vd with each v ∈ Vd being a tuple over a level schema Sl, and a family of "roll-up" relationships between the member subsets, defined by the dimension hierarchy.

A fact table F∈$\{F_1, ..., F_m\}$ of DM is composed of:

- The fact scheme S_f = { A_f , M_f }containing a finite, non-empty set of dimensional attributes A_f={ $A_{1,...}$ A_n} and a set of measure attributes M_f ={ $M_{1,...}$ M_n}. Each Ai ∈ A_f is linked with a level l ∈ L_D of the dimensions D, each Mj ∈ M_f with a τj . The domain of the attributes in A_f is defined as dom(Ai) = members(l) and dom(Mj) = dom(τj).
- The fact instance f(S_f), a set of tuples over {[dom(A_1) × ... × dom(A_n)], [dom(M_1) × ... × dom(M_m)]}. A tuple f ∈ f(S_f) is called a "cell" or "fact". Moreover, we call the values [f(A_1), ..., f(A_n)] the "coordinates" of a cell, modelling the multi-dimensional context for the measures [f(M_1), ..., f(M_m)]

Based on the formal definition of DM elements an analysis of the possible heterogeneities or conflict is required. Conflicts among independent and autonomous Data Marts may occur either at the schema or at the instance level and concern both dimensions and facts. Generally, heterogeneities among Data Marts covering the same application domain (e.g., sales figures of grocery stores) result from the use of either (1) different modelling patterns and methodologies, or (2) ambiguous domain vocabularies, or (3) a combination of both these factors. In recent years the heterogeneities that are specific to data warehouses and the multidimensional conceptual model are analyzed in [4], [5], [6]. A summary of the possible heterogeneities on dimension level are presented in the table 1. The facts logically depend on the dimension, thus all heterogeneities between corresponding dimensions must be repaired before the fact conflict. Based on that and for simplicity, we are considering only dimension conflict, leaving the fact conflicts aside.

Table 1. Dimension conflicts in multi-dimensional DM integration.

Conflict	Description	Formal definition
Naming conflicts	Different name labels for ontologi- cally same schema elements	name(D)≠name(D') name(l.K)≠name(l'.K) name(l.N)≠name(l'.N),
Non-corresponding Dimension schemas	Two dimension schemas do not have any common equivalent aggregation level	($LD \cap$ LD' = ø)
Inner-level corresponding:	Two dimension schemas have at least one, but not all equivalent level(s) in common and the base level are different	($LD \cap$ LD' = ø) and $l_0^D \neq l_0^{D'}$

Base-level corresponding	Two dimension schemas have at least one, but not all equivalent level(s) in common and the base level are equivalent	($LD \cap$ LD' = ø) and $l_0^D = l_0^{D\prime}$
Flat corresponding dimension schemas	Two dimension schemas have equivalent sets of aggregation levels with identical base levels, but the hierarchies do not match	LD = LD' && HD ≠ HD′
Inner level domain conflicts	Two attributes in level schemas of autonomous dimensions assign different sets of allowed values (i.e., different domains).	dom(l.K) ≠ dom(l'.K) и $l_0^D \neq l_0^D$
Base level domain conflict	Two attributes in level schemas of autonomous dimensions assign different sets of allowed values (i.e., different domains).	dom(l.K) ≠ dom(l'.K) и $l_0^D = l_0^D$
Non-dimensional attributes domain conflict:	Two attributes in level schemas of autonomous dimensions assign different sets of allowed values (i.e., different domains).	dom(*l.N*) ≠ dom(*l'.N)*

3. Formal algorithms for conflict detection

DWs and DMs are basically the implementation of multidimensional data models along with some tools for manipulating the multidimensional data. DMs implement/represent the multidimensional data instance either as relational systems (ROLAP), proprietary multidimensional systems (MOLAP), or a hybrid of both (HOLAP).

In ROLAP (relational online analytical processing) systems, the multidimensional data model is an E/R model and it is represented as a relational database (RDB) instance. Although ROLAP uses a relational database to represent the MD instance, the database must be carefully designed for analytical processing.

Star and Snowflake are the two best-known relational schemas that are specifically designed for ROLAP MD databases. ROLAP maps operations on multidimensional data model to standard relational operations. An advantage of these systems is that they can be easily integrated into other relational information systems [11], [12]. Based on [12], ROLAPs are the most common implementation of the multidimensional instance in DWs. The rich infrastructure of RDBs enables using standard, common and established techniques [13]. Throughout this paper the realization of a DM as a relational database, based on a star schema model, is being considered.

The idea of the first algorithm (fig. 1.) is to connect to autonomous ROLAP Data Marts and retrieve local meta-data. To infer the logical schema, it analyzes the public keys and primary keys. The algorithm accepts as input relational star schema and at the end returns the multidimensional data model elements, which will be used for comparing

with other MDM elements in order to automatically detect the possible heterogeneities and conflicts.

By analyzing the public keys and primary key, we can easily identify which of the tables are dimensional and which are fact tables, but some of the MDM elements cannot be detected automatically. Therefore, in the method being proposed, a qualified user decides which of the relational attributes are dimensional and which are non-dimensional.

Algorithm 1. *Metadata extraction*

Input: relational star schema R= $\{E_{1,...}E_n\}$

Output: multidimensional data model MDM=$\{F_1 = (A_{f1}, M_{f1}, dom(A_1), dom(F_1))$, $f(S_f 1), \ldots, F_n = (A_{fn}, M_{fn}, dom(A_n), dom(F_n)\ f(S_f))$, $D_1 =$ (d(SD1), dom(l.K1), dom(l.N1), LD1, S(LD1), HD1), ... , D_n = (d(SD1), dom(l.K1), dom(l.N1), LDn, S(LDn), HDn)}

```
1:  for all E ∈ R do
    //Defining dimensions and fact tables
2:      if(E_i.getPKCount() = 1) then
3:          Add(E_i) as D to MDM
4:      else if(T_i.getFKCount() > 0) then
5:          Add(E_i) as F to MDM
6:      end if
7:      end if
8:  end for
    //Defining dimension schema
9:  for all D ∈ MDM do
10:     for all D.Attr ∈ D_i
11:         if(askUser(D_i.Attr_j) = dimAtrr) then
12:             Add(Attr_j) as l to LD_i and HD_i
13:             Add(l.K_i.dataType()) as dom(l.K_i) to LD_i
14:         else
15:             Add(Attr_j) as l.N_i to Sl
16:             Add(l.N_i.dataType()) as dom(l.N_i) to Sl
17:         end if
18:     end for
19: end for
    //Defining dimensional and measure attributes in fact tables
20: for all F ∈ MDM do
21:     for all F.Attr ∈ F
22:         if(F_i.Attr_j = FK) then
23:             Add(Attr_j) as A_i to F_i
24:             Add(members(getLevel(Attr_j))) as dom(A_i) to F_i
25:         end if
26:         else
27:             Add(Attr_j) as M_i to F_i
28:             Add(M_i.dataType()) as dom(M_i) to F_i
29:         end if
30:     end for
31: end for
32: return MDM
```

Figure 1. Algorithm for metadata extraction

Based on the extracted multidimensional schema elements the second algorithm (fig. 2.) for conflict on dimension level is proposed. As input this algorithm accepts two dimensions with all of their constructing elements, and it returns a list with identified heterogeneities.

The basic idea here is to check automatically if these two dimensions describe the same real world entity, by introducing the rule – if two dimensions have at least one common or corresponding level in their schemas, they are corresponding and can be further analyzed for heterogeneities.

A new formula for calculating similarity between two dimensional attributes is introduced. The attributes' similarity is obtained from the similarity of their names (ds) and data type compatibility. The Levenshtein distance [15] is used as an algorithm for calculation of the name similarity. The notion of data type compatibility, in which attributes sharing their data type are more related than those that do not, has been assumed. The data types have been reduced to four basic ones: numeric, string, datetime and boolean.

If the inputted dimensions are not corresponding the algorithm stops and goes to the end. If the inputted dimensions are corresponding the algorithm goes through all of the dimensions elements and checks for conflicts and adds the detected conflict to a list.

The resulting list of all possible conflict between two corresponding dimensions can be used as a starting point of a data mart integration project.

Algorithm 2. Dimension level conflicts detection

Input: Dimensions D и D'
Output: List L of all conflicts between inputted dimensions

```
1:  for all l ∈ L_D do
2:          for all l' ∈ L_D' do
      //Calculating ds – degree of similarity between two levels
3:          ds=1-LevenshteinDistance(name(l_i),name(l'_j))/count_charact
    (name(l_i),name(l'_j))
4:                  if(dataType(l_i) = dataType(l_i) & ds > 0.6) then
5:                  l_i ≈ l'_j
6:                  count++
7:                  else
8:                  noncorres++
9:                  end if
10:         end for
11: end for
      //Defining corresponding dimensions
12: if(count=0 & (1 − LevenshteinDistance (name(D), name(D'))/
    count_charact (name(D),name(D')) < 0,5) then
13: Add NonCorspondDimSchemasConf(D,D') to L
14: go to end
15: else
16:         if(name(D) ≠ name(D')) then
      //Detecting and adding naming conflicts
17:         Add DimNameConf(name(D), name(D')) to L
18:         end if
19:         for all l ∈ L_D do
20:                 for all l' ∈ L_D' do
21:                         if(l_i ≈ l'_j & (name(l_i) ≠ name(l'_j))) then
22:                         Add LevelNameConf(name(l_i), name(l'_j)) to L
23:                         end if
    //Detecting and adding domain conflicts
24:                         if(l_i ≈ l'_j & dom(l_i.k) ≠ dom(l'_j.k')) then
25:                         Add DimDomainConf(dom(l_i.k), dom(l'_j.k')) to L
26:                         end if
27:                         for all l_i.N ∈ S_i do
28:                                  for all l'_j.N' ∈ S'_i do
29:                                 if(l_i ≈ l'_j & dom(l_i.N_k) ≠ dom(l'_j.N'_m)) then
30:                         Add NonDimDomainConf(dom(l_i.N_k), dom(l'_j.N'_m)) to L
31:                                         end if
32:                                 if(name(l_i.N_k)≠name(l'_j.N'_m)) then
33:                                 Add NonDimmNameConf(name(l_i.N_k),
                                                        name(l'_j.N'_m)) to L
34:                                 end if
35:                                 end for
36:                         end for
37:                 end for
38:         end for
39:         if(l_0^D ≠ l_0^D' & count < noncorres) then
40:         Add InnerSchemaCorresConf(S_D, S_D') to L
41:         else if(l_0^D = l_0^D' & count < noncorres)
42:         Add BaselevelCorresConf(S_D, S_D') to L
43:         if(H_D ≠ H_D') then
44:         Add DiffHierrarchyConf(H_D, H_D') to L
45:         end if
46:         for all v ∈ V_d do
47:                 for all v' ∈ V_d' do
48:                         if(v(1.K) = v'(1'.K) & r_D^{1→k} ≠ r_{D'}^{1'→k}) then
49:                         Add HeterRollUpFuncConf(v, v') to L
50:                         end if
51:                         for all v(1.N) ∈ V_d do
52:                                  for all v'(1'. N') ∈ V_d' do
53:                                     if(v(1.N_i) = v'(1'. N_i) & v(1. N_i) ≠ v'(1'. N_i)) then
54:                                 Add NonDimValueConf(v(l_i. N_i), v(l'_j. N'_j)) to L
55:                                         end if
56:                                 end for
57:                         end for
58:                 end for
59:         end for
60: end if
61: return L
```

Figure 2. Algorithm for dimension level conflict detection

4. Conclusion

This paper presents an approach to automating the conflicts detection process for heterogeneous data marts in order to shorten the integration process. The problem of integration of Data Marts has been observed, mainly focusing on conflicts occurring between semantically related dimensions. In particular, the focus has been on dimension schema conflicts, and their automatic detection.

To this end two algorithms for conflict detection are being proposed. The goal of the metadata extraction algorithm is to convert the relational DMs schema to multidimensional schema. This is done by connecting and extracting local metadata from the relational database. Currently the primary and foreign keys are being analyzed in order to detect dimensions and fact tables, and with so-called user metadata the rest of the multidimensional schema elements are being defined.

Based on the result of the first algorithm, the second algorithm aims to detect heterogeneities on dimension level. The novelty approach of this algorithm is the introduction of the correspondence between two dimensions, determined by calculation of the so called degree of similarity coefficient and data type capability.

References

1. Inmon, W. (2005). Building the Data Warehouse. John Wiley & Sons, New York, 4th edition.
2. Kimball, R. (2002). The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Datawarehouses. John Wiley & Sons, New York, 2nd edition.
3. Vassiliadis, P. and Sellis, T. K.: A survey of logical models for OLAP databases. (1999)
4. Tseng, F. S. C., Chen, C-W.: Integrating heterogeneous data warehouses using XML technologies. Journal of Information Science, 31(3):209-229 (2005)
5. Torlone, R.: Two approaches to the integration of heterogeneous data warehouses, Distrib. Parallel Databases 23(1):69-97 (2008)
6. Berger, S. and Schref, M. (2008). From federated databases to a federated data warehouse system. HICSS, 0:394.
7. Golfarelli, M., Maio, D., and Rizzi, S. (1998). The dimensional fact model: A conceptual model for data warehouses. Int. J. Cooperative Inf. Syst., 7(2-3):215-247.
8. Lazarova.V. 2012. Multitenancy in Cloud Computing. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education, October 5-6 th, 2012, Sofia, UNWE
9. P. Vassiliadis and T. K. Sellis. A Survey of Logical Models for OLAP Databases. SIGMOD Record, 28(4): pp. 64–69, 1999.
10. Diamantini, C., Potena, D., Sterfanov, G.: Two approaches for resolving dimension schema conflicts(2012) In: UNITE 2nd Doctoral Symposium, 11-12 October 2012, Sofia, Bulgaria
11. M. Jarke, M. Lenzerini, Y. Vassiliou, and P. Vassiliadis. Fundamentals of Data Warehouses. Springer, 1997.
12. R. Kimball and M. Ross. The Data Warehouse Toolkit: the Complete Guide to Dimensional Modeling. Wiley, 2nd edition, April 2002.
13. J. Wijsen. Condensed Representation of Database Repairs for Consistent Query Answering. In Proceedings of the International Conference on Database Theory (ICDT03), volume 2572 of Lecture Notes in Computer Science, pages 375–390. Springer, 2003.
14. Marzovanova, M., Advantages of using a system for intelligent tagging and search in unstructured data. In: International Conference on Application of Information and Communication Technology and Statistics in Economy and Education, October 5-6 th, 2012, Sofia, UNWE
15. В.И. Левенштейн (1965). "Двоичные коды с исправлением выпадений, вставок и замещений символов". Доклады Академий Наук СССР 163 (4): 845–8.

Creating Modern Web Applications Based on ASP.NET Web Forms and ASP.NET MVC

Iliya Nedyalkov

UNWE, 1700 Sofia, Bulgaria
iliya@idnc.eu

Abstract. Microsoft presented their vision for the web application development 10 years ago with the new ASP.NET platform. The pattern was Web Forms and this was then considered as a revolutionary approach for developing web applications - no spaghetti code, but code behind, rapid application development; opportunity of debugging; rich/extensive controls choice etc. There is a huge advantage in terms of usability in the Web Forms called View State. It genuinely maintains the state of the controls. Its drawback though is that it makes the web page heavier, hence in the era of fast web application is not really an option and applicable. Microsoft then presented to developers ASP.NET MVC pattern which works without View State and therefore applications based on MVC are lighter and faster than the Web Forms ones. MVC becomes more and more popular and a lot of Software architects withold from using Web Forms. We believe that both: Web Forms and MVC have their pros and cons and we will explain our thesis as to how these should be used simultaneously in one application to make the best of their advantages.

Keywords. Web Development, ASP.NET, Web Forms, MVC

1. Introduction

Microsoft presented their vision about development in 2003 with their multi-lingual .NET Framework. The part of it which is used for creating web applications is called ASP.NET Framework. In the past ASP.NET Framework was accompanying the Web Forms pattern. It is considered as a real revolution – it allowed the Desktop developers to start creating web applications in weeks. This has been really crucial since the development was in a position of having Desktop developers rather than such for Web.

Web Forms allows event driven development, separates back end source code from the HTML, and it uses ViewState to deal with the stateless HTTP and gives users with huge amount of Server Side controls.

Just like every technology Web Forms had some drawbacks. The utmost crucial one was the speed of execution. This is when Microsoft presented the MVC pattern which allows developers to create faster applications. A lot of programmers left behind the Web Forms and in favor of the MVC ones. There've been extensive discussions through the years whether both there's actual need of patterns. Both patterns are useful in different scenarios and they can even be applied simultaneously. Within this paper will be presented the current pros and cons of each pattern as well as a proposal as to how to use them both in the development of an application.

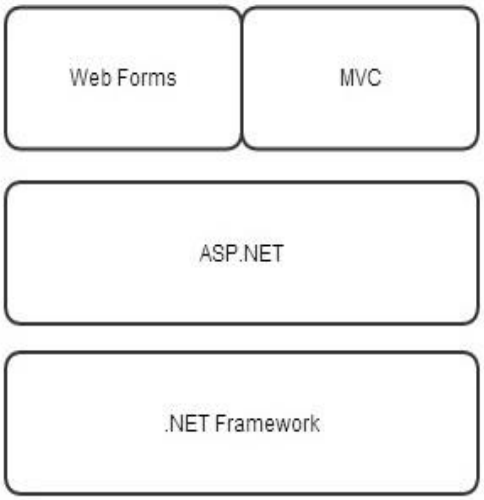


Fig.1 Microsoft vision of creating web applications

2. Exposure

Beyond the shadow of a doubt ASP.NET framework is nowadays one of the big players in creating modern web applications. There are ASP.NET developers who'd rather use Web Forms others would go with MVC. Using the experience gained through practice will be expressed the advantages and disadvantages of both development patterns as well as will be presented examples for which a particular way of development would be more suitable.

3. Web Forms

In the very beginning ASP.NET has been accompanying the Web Forms pattern. It's been considered as a breakthrough in the field of web development since it separates HTML and the source– no more spaghetti code. This allows HTML builders and Back end developers to work simultaneously.

Web Forms presents the idea of Event driven development. There are plenty of Rich Server Side controls which allow developers to drag and drop them on the development surface and to write source codes in the event handlers cause by the interaction of the users and the application. It was as simple as this to start developing Web Forms application.

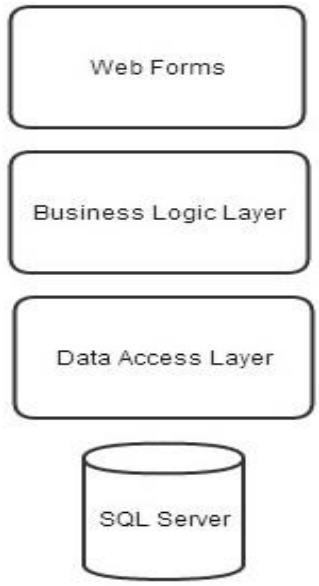
Rapid Application Development (RAD) is another big advantage of the web forms. The Rich Server Side controls allow just with drag and drop on the surface to create the visual part of the application. The raw HTML of a page is created in seconds. Most of the modern web applications store data in the Database and display that data in the browser. Speaking of which the Data binding takes place on every single control of the page which displays data from the Database. There are comprehensive controls which work with tabular

data like Grid View which have pagination and sorting functionalities and can be used in Web Forms applications.

Every web applications works identically. The web browser sends request to the web server, the web server does some calculations/executions and returns response the web browser. This communication is based on HTTP which is stateless protocol. This means that when there is a post back the controls of the page don't have state and they become empty. Web forms resolve this problem with View state. Using Web forms the ASP.NET developers don't think about the state of controls. They have state fullness out of the box.

When it comes to Web Forms the more it gives, the more it takes away. View state which helps developers dealing with stateless HTTP put extra kilobytes in the page. The more information the page contains the bigger the view state is. Web server controls are perfect for RAD but to be used with no alternations. They have difficulties in client scripts integrations and one could never guess how they would be rendered by the browser. Modern web applications have comprehensive Java Script codes and libraries and the importance of the HTML is increasing by the hour. In the era when everybody wants fast web applications the speed is crucial.

Nowadays the business demands not only fancy and fast applications but also excellent Search Engine Optimization (SEO). No doubt the Web Forms URLs are SEO unfriendly. The URLs will look like this http://www.XXXX.com/ProductDetails.aspx ?id=238



Pros

- No Spaghetti code
- RAD
- Data binding capabilities
- State fullness based on View State
- Event driven development

Cons

- Low performance
- View State
- Less control over HTML
- SEO unfriendly URLs

Fig 2 Pros and cons of Web Forms application

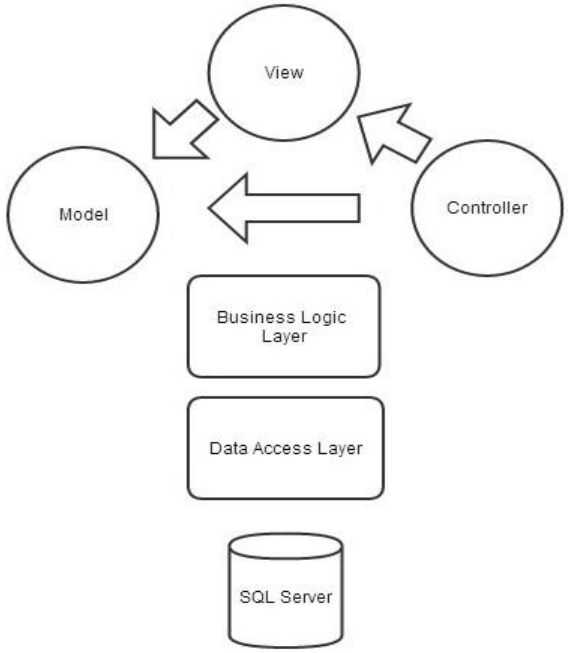
The Web Forms pattern is suitable for corporate information systems where there's need the HTML to be fancy, there're lots of data representations coming from the Database where Server side controls can take place and the speed of the applications is not crucial.

4. MVC

The drawbacks listed above forced Microsoft to move forward. They presented Model View Controller (MVC) pattern as a way to resolve Web Forms shortcomings. Based on the practical experience we think that MVC really resolved some of the Web Forms problems but it still has some disadvantages. We are going to propose cases in which choosing MVC would be the better solution than Web forms.

Model is considered as smart and could handle business rules, logic and data manipulation. It is a separate part of the controller and the view in MVC pattern. Controller receives requests, handles user interactions and input. It selects the right view in which to display the requested data. View presents the model data.

MVC performs better than Web Forms because there is no View state. Developers need to take care of the state management themselves. So there is not any kind of magic under the hood. Developers take better control on the application development.



Pros

- Performance (No View State)
- Full control over HTML – no server controls
- Support parallel development
- SEO friendly URLs

Cons

- More learning effort
- Slower development (Missing ready for use Server controls)

Fig 3 Pros and cons of MVC application

It is a common believe that the development of MVC applications is slower than development of Web Forms ones because there aren't rich server side controls out of the box. As explained above they make development pretty fast and straight forward. There is no

free lunch in MVC but this allows developers to have full control over HTML. They know that there won't be any surprises in rendering the HTML by the browser.

Parallel development has been very popular in the recent years. MVC allows three developers to work in Model/View/Controller simultaneously. This is pretty big advantage when the business is in a rush and there's a tight deadline.

MVC is comes with rich routing features and URLs are SEO friendly. They look like this this http://www.XXXX.com/ProductDetails/LongRedDress .

MVC is very suitable for presentational sites which will have a lot of simultaneous users. Avoiding View state the site will be lighter and will work faster. It is also suitable for web applications when their commercial part is bigger than the Back office. If the Back office is robust, development without Server side controls will not take long.

5. Web Forms and MVC collaboration

How to create modern web application? Use Web Forms or MVC? To create modern web application the developer may need both patterns in collaboration.

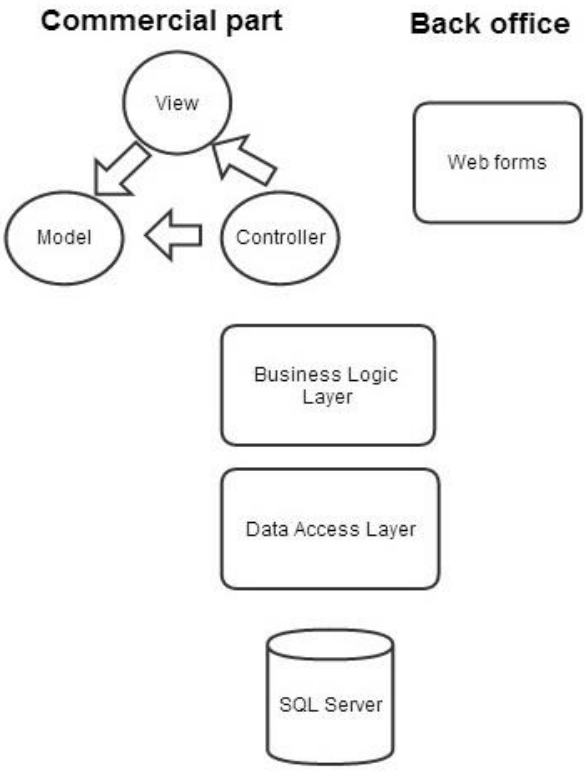


Fig 4 Web Forms and MVC in collaboration

Almost every web application has two parts – commercial one and back office. They may have the same number of pages. In a situation like this when it is not clear which pattern

will add more value in the development it would be recommended to use MVC and Web forms in a collaboration.

In the Back office it is recommended to use Web forms, most of the time Back office users are not that many like commercial ones and the speed of execution is not crucial. So we can use Web forms paradigm to create the Back office of the web applications and to take all the advantages of RAD.

In the commercial part where the speed is very important we propose using MVC. Commercial part is not using that standard server controls and we can avoid advantages like data binding and View state.

6. Conclusion

MVC pattern will not replace Web Forms. There are different cases in which the advantages of both ways of development could be applied. Even more there are cases in which they can be used in the same web application. Our proposal is to analyze the customer request very carefully before making a decision whether to use Web Forms, MVC or both of them in collaboration.

References

1. Marla Sukesh, *Web Forms vs. MVC*, (http://www.codeproject.com/Articles/528117/WebForms-vs-MVC)
2. *My View on ASP.NET Web Forms versus MVC* (http://weblogs.asp.net/ricardoperes/archive/2013/06/28/my-view-on-asp-net-web-forms-versus-mvc.aspx)
3. http://www.asp.net/web-forms
4. http://www.asp.net/mvc
5. http://msdn.microsoft.com/en-us/magazine/dd942833.aspx

Analysis of the Documents Processing Systems in Municipalities of Bulgaria

INA LYUBENOVA
University of National and World Economy - Sofia,
Department 'Information Technologies and Communications',
ina.lyubenova@gmail.com

Abstract:

The report examines the information systems in the municipalities of Bulgaria and in particular their systems for processing documents as one of the most common used. To determine the current status of information systems in the municipalities of Bulgaria a questionnaire has been developed and an empirical study has been conducted. I used part of this study to make an analysis of the local systems for processing documents. And after that I have analyzed how the availability of computers affects at the level of processing documents in municipalities. I used the resources of the statistical survey of the relationships to prove the hypotheses set out in the report. Based on the analysis, conclusions are made for improving municipal systems for processing documents, so the local administration could use it in their work.

Keywords: Municipality, local administration, systems for processing documents

The main objective of the business intelligent systems (BIS) is to facilitate the process of decision making in organizations, providing reliable basis for effective management. BIS use technology, software applications and practices for the collection, integration, analysis and presentation of business information necessary for decision making. The instruments implemented by BIS cover traditional forms of request, response and online analytical processing, as well as mining of patterns from data tables, maps with balanced performance and provide with adequate visualization.

This paper focuses on the business intelligent systems of the municipalities in Bulgaria. For the determination of their features a questionnaire is designed and an empirical research is conducted. The study was conducted on a representative sample of 54 municipalities in Bulgaria. Figure 1 shows the proportion of municipalities in the sample to all municipalities in Bulgaria. Studied municipalities accounted for approximately 22% of the population. According to many statistical sources, samples with a volume of more than 30 units are defined as larger samples and their results should be considered reliable in terms of the population of all 264 municipalities in Bulgaria.

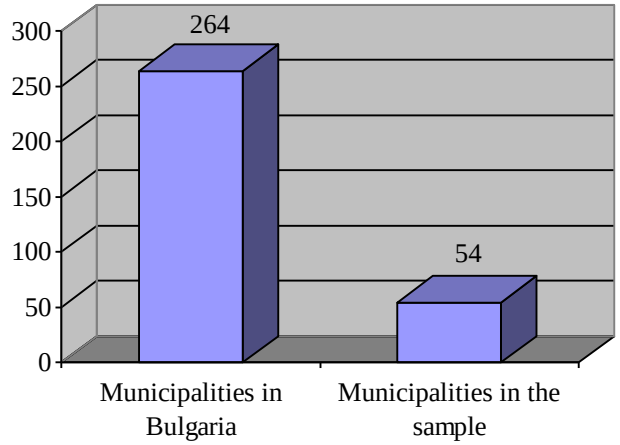


Figure 1. Municipalities in the sample compare to all the municipalities in Bulgaria

All the municipalities in the sample of the present study were selected randomly through a lottery selection and a function that generates random numbers. For the purposes of the study a stratified sample was used, and therefore the sample consists of municipalities from all 28 regions of Bulgaria. In this sense, referring to an administrative district in the Republic of Bulgaria, the possibility of each municipality to be in the sample is equal compared to other municipalities in this district. The way in which the sample is formed, gives us reason to believe the data obtained from the questionnaires is representative and we could treat the statistical results from the sample to the entire population of the municipalities in Bulgaria. Table 1 shows a list of the random chosen municipalities in the sample with their population according to the National Statistical Institute (NSI) 31.12.2010. The total number of residents in the surveyed municipalities is 3,418,906, which is almost 50% of the total population of the Republic Bulgaria.

DISTRICT	MUNICIPALITY	POPULATION
Blagoevgrad	Blagoevgrad	76812
	Petrich	55408
Burgas	Burag	206700
	Tzarevo	9411
Varna	Aksakovo	21972
	Dolni Chiflik	19141
	Dalgopol	14204
Veliko Tarnovo	Gorna Oriahovitza	48049

	Zlataritza	4584
	Svishtov	48351
Vidin	Vidin	64989
Vratza	Vratza	74648
	Oriahovo	12069
Gabrovo	Gabrovo	66321
	Drianovo	10315
	Triavna	12094
Dobrich	Kavarna	15657
	Krushari	5118
Kardgali	Ardino	12282
	Kardgali	68406
Kiustendil	Kocherinovo	5811
	Kiustendil	61944
Lovech	Lovech	52308
	Yablanitza	6334
Montana	Berkovitza	19256
Pazardgik	Pazardgik	120422
	Rakitovo	15418
Pernik	Breznik	7506
Pleven	Belene	10671
	Dolni Dabnik	14230
Plovdiv	Karlovo	53656
	Plovdiv	347611
	Hisaria	12768
Razgrad	Razgrad	53918
	Zavet	11110
Ruse	Borovo	6511
Silistra	Dulovo	28634
	Tutrakan	16637
Sliven	Nova Zagora	41199
	Tvarditza	14180
Smolian	Devin	12963
	Zlatograd	12112
Sofia (capital)	Stolichna	1259446
Sofia	Botevgrad	33529
	Kostinbrod	17448
Stara Zagora	Bratia Daskalovi	9625
	Kazanlak	75509

Targovishte	Antonovo	6395
	Targovishte	59855
Haskovo	Svilengrad	23867
	Dimitrovgrad	55869
Shumen	Varbitza	10318
Yambol	Elhovo	16332
	Yambol	75742
Total		**3418906**

Table 1. Municipalities in the sample and their population, NSI (www.nsi.bg)

The questionnaire was divided into several main sections, each of which consists of a different number of questions. Total number of questions in the questionnaire was 50. All questions in the survey are intended to clarify the current level of use of information technology in local government. Sections in the questionnaire are:

1. Employees in Information Technology (9 questions);
2. Hardware (7 questions);
3. Software (7 questions);
4. Databases (14 questions);
5. Website, online services (7 questions);
6. Business intelligent systems, early warning systems (6 questions);

The answers of question 6.2 presented in Figure 3. give us an idea of the use of information systems in the municipalities in Bulgaria. The study shows that a very high percentage of the municipalities have a documents processing system, over 80% or 47 of all 54 of the surveyed municipalities.

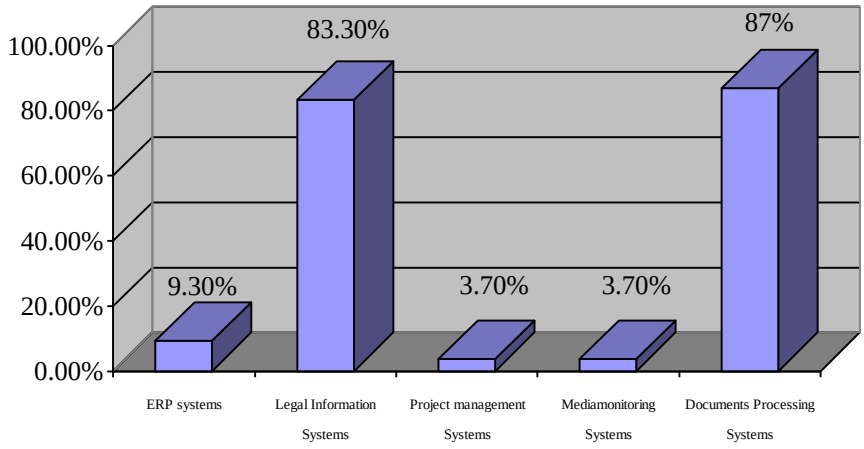


Figure 2. One dimensional frequency distribution of Question 6.2. Which of the following systems do you have in your municipality?

To analyze the documents processing systems in municipalities in Bulgaria, we will examine the relationship between the answers to the following questions from the questionnaire:

- Question 2.3. What percentage of computerized workstations do you have in your municipality?;
- Question 6.2.5 Do you have a documents processing system in your municipality?;

Next we should analyze the statement is there a statistical correlation between the availability of a documents processing system and the extent of computerization of workplaces in municipalities. For this purpose we will use the Chi-square method to study the relationship, as the phenomena are located on the weak measurement scales. In Table 2 is presented the raw data for the study.

Table 2

f_{ij} Computerized workstations	Available documents processing system		Total
	yes	no	
Up to 50% of the workstations	1	2	3
From 51% to 75% of the workstations	4	3	7
Above 75% of the workstations	40	4	44
Total	45	9	54

We define zero (H_0) and alternative (H_1) hypothesis. The zero hypothesis would be the assumption that between the availability of a documents processing system and the extent of computerization there is no objective connection. The alternative hypothesis is the assumption that between the two phenomena studied there is an objective, not accidental connection.

We fix error risk of first kind (α), then define risk error of first kind α =0.05.

The formula by which we evaluate the empirical characteristic has the form:

$$\chi^2_{em} = \sum_i \sum_j \frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}}$$

From here we find the theoretical frequencies shown in Table 3 using the formula:

$$\hat{f}_{ij} = \frac{\sum_i f_{\bullet i} \sum_j f_{\bullet j}}{\sum_i \sum_j f_{ij}}$$

Table 3. Theoretical frequencies $\hat{f}_{ij}$

$\hat{f}_{ij}$ Computerized workstations	Available documents processing system		Total
	yes	no	
Up to 50% of the workstations	2,5	0,5	3
From 51% to 75% of the workstations	5,83	1,17	7
Above 75% of the workstations	36,67	7,33	44
Total	45	9	54

Then we find the differences $\left(f_{ij} - \hat{f}_{ij}\right)$ and they are squared (Table 4):

Table 4.

$\left(f_{ij} - \hat{f}_{ij}\right)^2$ Computerized workstations	Available documents processing system	
	yes	no
Up to 50% of the workstations	2,25	2,25
From 51% to 75% of the workstations	3,35	3,35
Above 75% of the workstations	11,09	11,09

Finally we find the ratio $\frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}}$ shown in Table 5:

Table 5.

$\frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}}$	Available documents processing system		Total
	yes	no	

Computerized workstations			
Up to 50% of the workstations	0,9	4,5	5,4
From 51% to 75% of the workstations	0,57	2,86	3,43
Above 75% of the workstations	0,3	1,51	1,81
Total	1,77	8,87	10,64

The number calculated in the last cell of the table is χ^2_{em} *– the empirical feature of the hypothesis.*

χ^2_{em} *=10,64*

We determine the theoretical feature of the hypothesis – χ^2_t *. When parameters*

α *=0,05 and* f *= (p-1)(k-1) = (3-1)(2-1) = 2:*

(χ^2_t *=0.05,* α *=2) =* f *5,99*

From here we see that $\chi^2_{em} > \chi^2_t$ *=> There is a statistical correlation between the availability of a documents processing system and the level of computerization in the municipalities.*

To evaluate the relationship we will use ratios to assess the narrowness of the already proven relationship. We will present two coefficients that are most appropriate for this case.

Cramer's V coefficient:

$$V^2 = \frac{\chi^2_{em}}{\sum_i \sum_j f_{ij}[\min(p-1), \min(k-1)]} = \frac{10,64}{54(2-1)} = 0,197$$

Pearson coefficient:

$$C = \sqrt{\frac{\chi^2_{em}}{\chi^2_{em} + \sum_i \sum_j f_{ij}}} = \sqrt{\frac{10{,}64}{10{,}64 + 54}} = 0{,}165$$

The analysis has shown that there is a statistically significant relationship between the availability of a documents processing system and the extent of computerization of workstations in the municipalities in Bulgaria, but the relationship is weak. According to the Cramer's coefficient (= 0,197), only about 19.7% of the differences in the extent of computerization of workstations in municipalities depend on the availability of documents processing system.

This study aims to examine the problem of providing the municipalities with information technology and in particular with documents processing systems and to establish whether there is any statistical dependence associated with the presence of a sufficient number of computerized workstations in the municipalities of Bulgaria and to contribute to improvement in the efficiency of public services that municipalities provide through implementation and improvement of the current hardware and software solutions.

References:

1.www.nsi.bg

2.Съйкова И., Стойкова А, Съйкова С., Статистическо изследване на зависимости, УИ Стопанство, С., 2002

Evaluation of Applicability of the Biometric Technology in the Internet

Mihail Konchev

UNWE, 1700 Sofia, Bulgaria
mivailov@abv.bg

Abstract. Identity theft is one of the problems facing us today. It is defined as the pursuit of illegal activities with personal data for financial abuse. Development of information technology enables the use of security systems based on the recognition of unique physical characteristics. This suggests the introduction of new authentication systems based on biometric authentication.

The purposes of the paper are:

- Summarize some negative consequences arising from online anonymity;
- To present the traditional approaches used in the process of authentication on the web;
- To propose a model for evaluation of biometric technologies, in order to determine their degree of relevance to the needs of authentication and identification on the Interne
- To make a comparative analysis of traditional approaches and biometric technologies.

Keywords: biometric technology, voice recognition, web security

1. Introduction

Security is the avoidance of risk or danger [1]. The safety is a basic need for people related to different spheres of life, including the use of Internet. The most critical elements of internet security are web servers and applications that interact with multiple users with a view to share data, financial transactions, online banking, providing web services and more. In support of this are the results of a research of security vendor Cenzic "First and foremost, 99% of all applications tested in 2012 have one or more serious security vulnerabilities. And with a median number of vulnerabilities per app of 13, it's no wonder that application-level attacks are a focus for hackers."[2].

The essence of the Internet, in terms of its security, is the subject of increasing interest in the theory and analyzed by a number of authors. The Internet has no centralized governance in either technological implementation or policies for access and usage; each constituent network sets its own policies [3,39].

Analyzing this definition of the essence of the Internet, from the security point of view, the lack of specific rules for access and authentication of users in the network is a prerequisite for performing a number of attacks aimed to compromising the security of the web applications.

In support of this claim, the research's results of Cenzic show that: "At 26% of the total, Cross Site Scripting (XSS) was the most frequently found vulnerability in apps tested in 2012. Information Leakage and Session Management Errors follow in frequency, each at 16% of total vulnerabilities found. Authentication and Authorization (13%), Cross Site Request Forgery (CSRF) (8%) SQL Injection (6%), Web Server Version (5%), Remote Code Execution (5%), Web Server Configuration (3%), and Unauthorized Directory Access (2%) round out the 2012 vulnerability population"[2].

"Hacker" is a term lasting entered the IT terminology. It is used for persons who illegally penetrate foreign systems or inflict great harm, using computers as a weapon. However, among the experts are dealing with writing code that term means to find a quick and easy solution to a problem or a clever way to do something [4]. In this report the term "hacker" will be perceived as a synonym of "ineligible person" or "intruder". In specialized publications on the topic of security in the global network, most authors adopt a uniform qualification for the perpetrators of the attacks, namely: "unauthenticated user", "ineligible person", "intruder". Such an approach takes and Shiflett.Ch. Striking is the definition given by the author about the process of authentication, which is not quite accurate in all its details, "Authentication is the process by which prove the user's identity ..." [6]. The author defines authentication, a process that proves the identity of the user, which is true, but incomplete. An user can access someone else's data to a third person, which makes it really illegitimate user, but not for the process of authentication, which defines it as legitimate. A more precise definition is given by the team of Apache, according to which: "Authentication is any process by which you verify that someone is who they claim they are" [5]. This definition defines more precisely the process of authentication, but leaves open the question about the actual identity of the user.

Summarizing the presented views, we can define identity as one of the biggest problems on the Internet, from a security perspective.

2. The problem of the Internet anonymity

The growth of the Internet in the context of e-commerce, social networking and communication tools is interest to researchers from different scientific fields (communication, social psychology, computer security, etc.). They set out to explore the different behaviors in this type of social environment. "Computer mediated communication" (CMC) enters IT terminology and definition for mode of communication in this environment. [7] Particular interest to researchers is the growing importance of anonymity and its effects in CMC, as its determinant.

In the context of Internet communication (e-mail, newsgroup article, web page, pamphlet, book, rumour, etc.) the term "anonymity" is found most often in three varieties: anonymity, pseudonymity and a variant of pseudonymity – deception. Anonymity means that the real author of some communication unit is not shown. It can be implemented to make it impossible or very difficult to find out the real author. A common variant of anonymity is pseudonymity, where another name than the real author is shown. The pseudonym is sometimes kept very secret, sometimes the real name behind a pseudonym is openly known [8]. A variant of pseudonymity is deception, where a person intentionally tries to give the impression of being someone else, or of having different authority or expertise [9,10].

The analysis of the problem of Internet anonymity, literature showed that it considered two main aspects - positive and negative implications. *Looking at the Internet, within the social context, as "no centralized governance in either technological implementation or policies for access and usage ... [3]" it can be considered as a "deindividuating" environment, making conditional analogue to the definition of crowd in the "Social theories on anonymity".*

Social scientists have focused on the negative aspects of the anonymity for long time until now. Research in this direction has highlighted the two different theories: Deindividuation theory and Social Identity Theory of Deindividuation. Deindividuation

theory asserts that being in a large group provides a degree of anonymity, which allows one to avoid responsibility for his actions - thus shaking off usual social controls and becoming more impulsive, irrational, aggressive and even violent [11,12,13]. But recent research argues for a reconceptualisation of deindividuation. It appears deindividuation is not a loss of individual identity, but may be better construed as a transition to a social identity - Social Identity Theory of Deindividuation. This theory uses the concept of social identification. The theory states that in a crowd, and other "deindividuating" settings, the person does not simply lose a sense of individuality, but makes a transition from a personal identity to a social identity [14].

Within the social context, cited theories agree in the opinion, that in the conditions of anonymity, individuals lose their sense of self or just shifts their sense of self, hence the person loss the sense of responsibility for the actions taken.

Consequences of internet anonymity is widely discussed topic in the scientific community. It is considered by experts in various scientific fields, emphasizing the positive and negative aspects. Many authors consider anonymity as a *constitutional right and right of privacy*. This aspect is very important on the internet, where personal data could be gained, stored, processed and abused very easily [17]. *Social benefits* is another aspect of anonymity on the Internet. Anonymity reduced likelihood of retaliation can encourage whistleblowers to draw attention to serious problem. It open public or private discussions of potentially embarrassing personal problem [17]. Cooper have proposed that easy access, affordability and anonymity of the internet intensifies and accelerates online sexual activity. Variations in technological variables such as control of internet access and internet skill level should explain a substantial portion of the variance in people's online sexual activity [18,19].

From the security point of view, the lack of specific rules for access and authentication of users in the global network and with a sense of lack of responsibility is a prerequisite for anonymity can be used to protect a criminal performing many different crimes such as slander, illegal threats, racial agitation, fraud, intentional damage such as distribution of computer viruses, hacker attacks, etc.

To some extent, the phrase "internet anonymity" can be seen as an oxymoron (a combination of antitheses). This claim comes from the answer to the question "is there an absolute anonymity on the Internet?". Technologically, one of the specifications of the TCP / IP protocol is how data should be addressed, described in RFC 1122 and 1123 [15,16]. This requirement presupposes identity between corresponding sides. In the Internet network it is the IP address (physical internet address). In essence, it uniquely identifies every internet computer and it is generally possible to trace each user. Here we should note the existence of exceptions to this rule - use of third-party computer as a redirector, though, those machines usually keep the logs anyway. Unique identification of each computer on the network via IP, suggesting that absolute anonymity on the Internet is an illusion. But the IP address uniquely identifies the physical internet address of the computer, not the user, using the computer. What would be the identification of a user when he is connected to the internet, using a temporary IP number in a free wireless network?! ***This allows us to maintain claim that anonymity on the Internet exists, in terms of the identification of the user identity.***

The identity of a person, place or thing is determined by its characteristics, features or circumstances by which a thing or person is definitely recognisable or known, it is the set of behavioural and personal characteristics by which an individual is recognizable as a member of a group [20].

According to Kollock and Smith "Online interaction strips away many of the clues and signs that are part of face-to-face interaction. This poverty of signals is both a limitation and a resource, making certain kinds of interaction more difficult but also providing room to play with one's identity" [21,22]. In their publications the authors examined from a technological point of view, different options for identification on the Internet, through indirect "clues." Email account name is generally perceived to refer to a single person. It may provide some contextual information about the writer, information that, while quite sketch, may be the only such cues – and even if not, it reveals he wants to stay anonymous. Email addresses from "official" domains are therefore generally taken as real names with all the consequences, since the author is usually employee/student/or some other person, whose identity was verified in the real word. And there is an accountability involved, too. The domain name gives contextual clues about the author – and about the reliability of the information. A work or school-based account name is known within the organization and there are many people who can make a direct connection between the e-mail address and the real-world person. On the other hand, freemail providers do not guarantee anything and offer wide range of anonymity. Another source of identity could be email signature. Email signature is text that is automatically attached to the bottom of email message. Protocols and services can be understood by the application/network the user uses (HTTP, FTP, VoIP) or specialized chat networks.

The analysis of these options indicates their low degree of reliability. In support of this, we should point out that according Donath.J. "Identity cues provided by users are not always reliable. The account name in the header of an email can be faked, identity claims can be false, social cues can be deliberately misleading" [9]. ***Summarizing the presented views, we can define identity as one of the biggest problems on the Internet, from a security perspective.***

The use of biometrics as a possible solution to the problem of identification is a relatively new approach. It is a combination of different technologies using biometric data, such as: iris, retina, finger prints, hand geometry, palms, voice, face, etc. A biometric is a physical characteristic, a measure of a biological trait. Biometrics has a very useful application in security; it can be used to authenticate a person's identity and control access to a restricted area or electronic system, based on the premise that certain of these physical characteristics can be used to uniquely identify individuals [23].

Electronic signature is any electronic means that indicates the person who claims to have written a message is the one who wrote it (and that the message received is the one that was sent). It also serves to verify your identity that's why you have to sign credit card receipts, after all. But in many of these cases, identity isn't really that big of a deal. In cases where it is major financial transactions, business operations, or legal filings a notary is required to ensure that you, the person signing the document, are actually who you claim to be [25, 26]

Logical questions arise: 1) The degree to which these technologies are applicable for the purposes of identification and authentication on the Internet?; 2) What are the advantages and disadvantages over traditional approaches?

One of the tasks of this report is to propose a model for evaluation of biometric technologies, in order to determine their degree of relevance to the needs of authentication and identification on the Internet.

3. Evaluation model of technologies as an authenticator in a networked environment

The scope of this model is limited to the selection of technologies as an authenticator in a networked environment. Biometrics as a physical access, e-commerce, and monitoring technology is beyond the scope of this model. ***In relation to the target is to be noted that the evaluation given by the model can be interpreted only under the condition that the model is unchanged.***

The purpose of the present model is:

1) To evaluate the technologies as an authenticator in a networked environment in qualitative and quantitative indicators, according to the following criteria:

   1. **Require additional hardware for implementation** – *from an architectural point of view* relevant to the assessment of biometric technology as an authenticator in a networked environment is the need of additional hardware for implementation;
   2. **Required software and technology for web implementation** – this criterion assesses the necessary software for implementation in a networked environment;
   3. **Opportunity for remote authentication** – this criterion refers to the ability of the evaluated technologies for remote authentication;
   4. **Hacking** – this criterion summarizes the most common hacker attacks to the technology;
   5. **Performance factors** - this criterion assesses the factors influencing the performance of the technology;
   6. **Level of accuracy** - this criterion assesses Crossover Error Rate (CER). In percentage terms at which point False Rejection Rate (FRR) = False Acceptance Rate (FAR). It is necessary to note that the CER vary from mechanism-to-mechanism and vendor-to-vendor. Technology assessment for this criterion is perceived average assessment rates;
   7. **Storage size** – this criterion assesses the size of information for access;
   8. **Storage device** - this criterion summarizes compatible with the technology of devices for storing data for access;
   9. **System decision time** - this criterion assesses the time for which the system returns answer;
   10. **Cost** – this criterion evaluates the cost of the license for technologies needed for development and implementation of the technology in a networked environment;

2) The final score for each criterion to be formed by a scoring system ranging from 0 to 100 points. Assessment is equivalent to the following scale (Table 1) as described below approach.

Points	Evaluation
from 0 to 25 points	Weak
from 26 to 50 points	Good
from 51 to 75 points	Better
from 76 to 100 points	Excellent

Table 1. Rating Scale

Assessment approach:

1. Final score for each criteria is formed as defined in the purpose of the rating scale model;
2. Criteria for evaluating the software by dichotomous question requires answering "Yes" or "No" are measured with 100 points for a positive response and 0 points for negative;
3. Evaluation requires a comparison with other technologies using proportionate approach. Under the approach quantified indicators for benchmarking technologies are proportional to the points given to them for evaluation. The technology which show the best result in each criterion are given scores of 100 points. The final grade is rounded to the nearest whole number.
4. The final mark is formed by calculating the average of collected criteria points;
5. The results obtained are set equal to the rating scale (Table 1 Rating Scale);
6. The results obtained make conclusions about the relevance of the technology as an authenticator in a networked environment.

4. Application of the proposed evaluation model

According to the objectives of the report, the proposed valuation model applies to the following technologies: 1) Retinal Scan; 2) Iris Scan; 3) Fingerprints; 4) Hand Geometry; 5) Signature Dynamics; 6) Voice Dynamics; 7) Single-factor authentication , 8) Multi-Factor authentication. Single-factor authentication is a technology witch traditional websites requiring just a user's password ask for something the user knows (a password) or ask for something the user physically has (an RFID card). But multi-factor authentication on the other hand requires verification from at least two of the different approaches and is considerably more secure. A more practical approach to multi-factor authentication is to implement options from the other two categories, such as requiring a password and a confirmation token sent to the user via phone either by SMS or voice call. The results are summarized in *Table 2*, *Table 3*, and graphically represented in *Figure 1*.

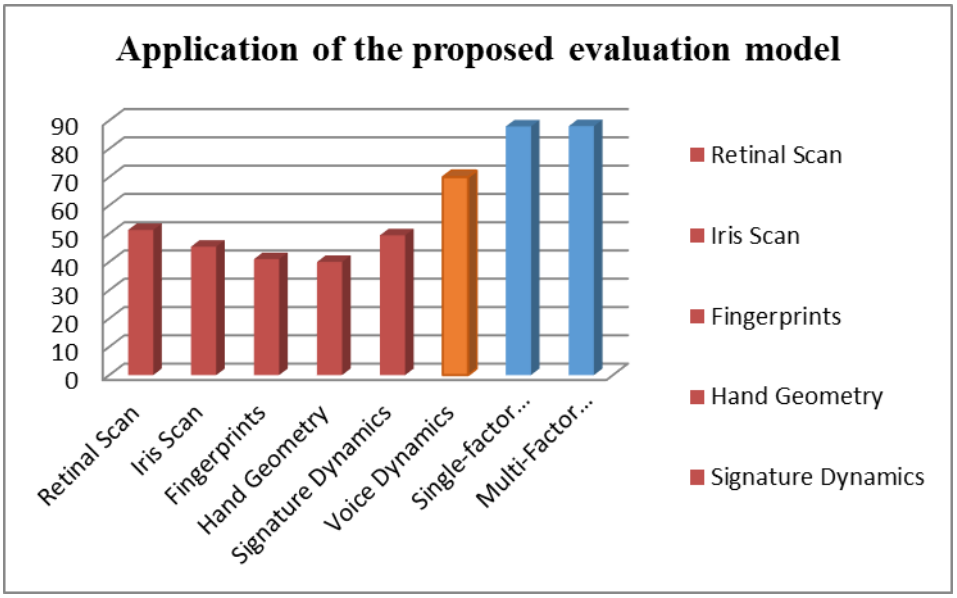


Figure 1. The aggregation results of application of the evaluation model

After application of the proposed evaluation model can draw the following conclusions for the evaluated technologies:

	Criteria									
Authentication by	Require additional hardware for implementation	Required software and Technology for web implementation	Opportunity for remote authentication	Hacking	Performance factors	Level of accuracy (Crossover Accuracy %)	Storage size	Storage device	System decision time	Cost

Hand Geometry	Yes	Graphical User Interface implemented with a high programming language as C or Java [29]	Yes	Artificial hand; Replay attack; Directly database compromise; Sniffing data; [32]	Hand injury; Age; Jewelry [24]	High .2% [23,24]	9 bytes – 3k bytes	smart cards; floppy disks; databases; cell phones. [31]		.NET Framework license; Java (GPL)
Fingerprints	Yes	Graphical User Interface implemented with a high programming language as C or Java [29]	Yes	Artificial fingers; Replay attack; Directly database compromise; Sniffing data; [32]	Dryness Dirt; Worn; Aged fingertips [24]	High .2% [23,24]	250 – 1000 bytes [27]	smart cards; floppy disks; databases; cell phones. [31]	~ 8 sec. (125 000	.NET Framework license; Java (GPL)
Iris Scan	Yes	Graphical User Interface implemented with a high programming language as C or Java [29]	Yes	Special contacts; Fake eyes; Replay attack; Directly database compromise; Sniffing data;[32]	Poor Lighting Movement [24]	Very High .000763% [23,24]	Iris Code® 512 bytes [28]	smart cards; floppy disks; databases; cell phones [31]	~1,7 sec [34]	.NET Framework license; Java (GPL)
Retinal Scan	Yes	Graphical User Interface implemented with a high programming language as C or Java [29]	Yes	Fake eyes; Replay attack; Directly database compromise; Sniffing data; [32]	Glasses Difficult to use; [24]	Very High.0000001% [23,24]	9 bytes – 3k bytes [24]	smart cards floppy disks databases; cell phones [31]	Over 1 min. [27]	.NET Framework license Java (GPL)

Signature Dynamics	Yes	Graphical User Interface implemented with a high programming language as C or Java [29]	Yes	Replay attack; Directly database compromise; Sniffing data; [32]	Changing or erratic Signatures [24]	High 2% [23,24]		smart cards; floppy disks; databases; cell phones. [31]		.NET Framework license; Java (GPL)
Voice Dynamics	No	Adobe Flash jQuery Server Script Language	Yes	Replay attack; Directly database compromise; Sniffing data; [32]	Noise; Colds; Weathe; Age; Equipment; Environment; Illness. [24]	High 2% [23,24]	1000 bytes [23]	smart cards; floppy disks; databases; cell phones. [31]	0,5 sec [30]	GPL
Single-factor authentication	No	Client-Server technology	Yes	**Authentication** Brute Force Insufficient Authentication Weak Password Recovery Validation **Authorization** Credential/Session Prediction Insufficient Authorization Insufficient Session Expiration Session Fixation [33]	**Theoretical attack** MD5 [37]	High 0%	**Internal state size** MD5 – 128 bits; SHA 0 – 160 bits; SHA 1 – 160 bits; SHA 2 – 256 bits. [36]	smart cards; floppy disks; databases; cell phones. [31]	ms-sec	

Multy-Factor authentication	No	Client-Server technology Cell phones;	Yes	Social engineering; Combination of single factor authentication hacking; [33]	NA	Very High 0%	**Internal state size** MD5 – 128 bits; SHA 0 – 160 bits; SHA 1 – 160 bits;SHA 2 – 256 bits. [36]		~ 2 min	

Table 2. The evaluation model matrix

	Criteria (points)										
Authentication by	Require additional hardware for implementation	Required software and Technology for web implementation	Opportunity for remote authentication	Hacking	Performance factors	Level of accuracy (Crossover Accuracy %)	Storage size	Storage device	System decision time	Cost	**AVERAGE**
Retinal Scan	0	0	100	50	50	100	12	100	1	100	**51,3**
Iris Scan	0	0	100	40	50	1	34	100	29	100	**45,4**
Fingerprints	0	0	100	50	25	1	28	100	6	100	**41**
Hand Geometry	0	0	100	50	33	1	12	100	4	100	**40**
Signature Dynamics	0	0	100	67	50	1	70	100	6	100	**49,4**
Voice Dynamics	100	100	100	67	14	1	18	100	100	100	**70**
Single-factor authentication	100	100	100	29	100	100	100	100	50	100	**87,9**
Multi-Factor authentication	100	80	100	100	100	100	100	100	0	100	**88**

Table 3. The results of application of the evaluation model

Empirical analysis shows that the Single-factor and Multi-factor authentication methods show excellent results under the proposed evaluation model. From the results it is clear that the most vulnerable to attacks in order to compromise the security are Single-factor authentication methods. By this standard, maximum results indicate Multi-factor authentication methods, but at the expense of high latency of the overall process of authentication. In connection with the subject of the report interest is the evaluation of biometric technologies in terms of their applicability to the Internet as a means of authentication. According to the evaluation model analyzed biometric technology show good

score on all criteria. Retinal scan and Voice dynamics, according to the same rating scale showed very good. The advantages of Retinal scan technology to Voice dynamics are few factors that affect her work and her great degree of accuracy, expressed by "Crossover Error Rate" in rates. With all other parameters Voice dynamics showed better results. Main advantages of Voice dynamics over other technologies are:

1) The lack of need for additional hardware, except a microphone on the use of technology;
2) Implementation of the technology in the Web environment can be realized through the use of free technologies such as jQuery library, Flash player and programming language to connect to the server as: PHP, Pirl and others. To implement the "Retinal scan"technology requires the use of high programming language as C or Java.

The analysis of the results of the application of the evaluation model shows that biometric technologies I can be administered alone as a means of authentication on the Internet, due to some weaknesses in their security as hacking and degree of accuracy, expressed by "Crossover Error Rate". They can be used as part of milti-factor authentication process using their main advantage over traditional approaches - their ability to identify the actual user, do not report outlined opportunities to discredit them. ***The main utility of this technology is to provide identification during the authentication process***.

In security there are 3 commonly referenced factors by which a user can authenticate [38]:

1) Something you know (ID, PIN, Password, Passphrase);
2) Something you have (Certificates, and tokens including One-Time-Passwords, smart cards, key cards, USB fobs, challenge/response);
3) Something you are (Biometrics).

5. Conclusion

"Best Security Practices" dictate that network authentication use minimally 2 and preferably 3 factors as part of a strong authentication scheme.

1. A voice biometric or "voice print," is as unique to an individual as a palm or finger print.
2. Any authentify application that employs a voice channel is able to add voice biometric authentication to the process for even higher levels of authentication and security.
3. Any process that uses biometrics can be tuned to raise or lower the probability level at which a new model is accepted or rejected when compared to the original. This permits to balance the two types of errors – the false positives and the false negatives.
4. The point, referred to "Equal Error Rate," at which, probabilistically speaking, there is an equal chance of being falsely rejected, or falsely accepted. Beyond the Equal Error Rate, the biometric can be tuned to allow more false positives or false negatives depending on the application's security requirements.

References

1. Shiflett.Ch., „Essential PHP Security", "ZeST Press Publishing House, Ltd", 2007, ISBN-13:978-954-9341-20-1 p.vii;
2. 2012 Trends Report: Application Security Risks". Cenzic, Inc. 11 March 2012. Retrieved 9 July 2012;
3. Internet., http://en.wikipedia.org/wiki/Internet;
4. Herrigton.J., "PHP HACKS, Tips & Tools for Creating Dynamic Web Sites", "ZeST Press Publishing House, Ltd", 2007, ISBN-13:978-954-9341-15-7 p. xiii;
5. Apache, Authentication and Authorization, http://httpd.apache.org/docs/2.2/howto/auth.html;
6. Shiflett.Ch., „Essential PHP Security", "ZeST Press Publishing House, Ltd", 2007, ISBN-13:978-954-9341-20-1 p.67;
7. McQuail D. (2005). Mcquail's Mass Communication Theory. 5th ed. London: SAGE Publications;
8. Palme.J., Berglund M. "Anonymity on the Internet", http://people.dsv.su.se/~jpalme/society/anonymity.pdf;
9. Donath. J "Identity and Deception in the Virtual Community by udith, in Kollock. P.",1996;
10. Smith M.," Communities in Cyberspace, Routledge, London"., 1999, http://smg.media.mit.edu/people/Judith/Identity/IdentityD eception.html;
11. Festinger, L., Pepitone, A., & Newcomb, T. (1952), Some consequences of de-individuation in a group. Journal of Abnormal and Social Psychology, 47: 382-389;
12. Zimbardo, P. G. (1969), The human choice - Individuation, reason, and order vs. deindividuation, impulse and chaos. In: W.J. Arnold & D. Levine (Eds.), Nebraska Symposium on Motivation (Vol. 17, pp. 237-307), Lincoln, NE: University of Nebraska Press;
13. Diener, E., Fraser, S.C., Beaman, A. L. & Kelem, R. T. (1976), Effects of deindividuating variables on stealing by Halloween trick-or-treaters. Journal of Personality and Social Psychology 33:178- 183;
14. Reicher, S. D. (1987), Crowd behaviour as social action. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), Rediscovering the social group: A self-categorization theory (pp. 171-202), Oxford, England: Basil Blackwell;
15. RFC 1122, Requirements for Internet Hosts – Communication Layers, R. Braden (ed.), October 1989;
16. RFC 1123, Requirements for Internet Hosts – Application and Support, R. Braden (ed.), October 1989;
17. Wallace, Jonathan, D. (1999) Nameless in Cyberspace – Anonymity on the Internet. Briefing Paper no. 54 Cato Institute, Washington DC, 1999, http://www.cato.org/pubs/briefs/bp-054es.html;
18. Cooper, A. (Ed.). (2002). Sex and the internet: A Guidebook for Clinicians. New York, NY: Brunner & Routledge;
19. Palme,J., "Anonymity on the Internet. Stockholm University", 2002, http://dsv.su.se/jpalme/society/anonymity.pdf;
20. American Heritage Dictionary of the English Language (2000), Fourth Edition, Houghton Mifflin Company – electronic version is available on the internet address www.dictionary.com;
21. Kollock, Peter, and Marc Smith, Managing the virtual commons: cooperation and conflict in computer communities. In Susan Herring (Ed.) Computer-Mediated Communication: Linguistic, Social, and CrossCultural Perspectives. Amsterdam: John Benjamins, 1996;
22. Kollock, Peter & Smith, Marc A., Communities in cyberspace. London: Routledge, 1999;
23. Myers.L., "An Exploration of Voice Biometrics", 2004, http://www.sans.org/reading room/whitepapers/authentication/exploration-voice-biometrics-1436;
24. Walker.S.," Biometric Selection: Body Parts Online", 2002, http://www.sans.org/reading room/whitepapers/authentication/biometric-selection-body-parts-online-139;

25. Null.Ch.,"E-signatures: The complete guide to paperless signing", 2013, http://www.pcworld.com/article/2035744/e-signatures-the-complete-guide-to-paperless-signing.html;
26. Electronic signature, http://en.wikipedia.org/wiki/Electronic_signature;
27. Spinella.E.," Biometric Scanning Technologies: Finger, Facial and Retinal Scanning", 2003, http://www.sans.org/reading-room/whitepapers/authentication/biometric-scanning-technologies-finger-facial-retinal-scanning-1177;
28. Khaw.P.," Iris Recognition Technology for Improved Authentication", 2002, http://www.sans.org/reading room/whitepapers/authentication/iris-recognition-technology-improved-authentication-132;
29. Iris Recognition System, http://www.n3trino.com/sites/default/files/Iris.pdf;
30. Schmidt, Regina. http://www.biometritech.com/features/090303nu.htm;
31. http://www.voice-security.com/NewVoice.html;
32. Beals.B., "Biometrics: Hack Proof?", 2002, http://www.giac.org/paper/gsec/2282/biometrics-hack-proof/103919;
33. Web Application Security Consortium: Threat Classification, 2004, http://projects.webappsec.org/f/WASC-TC-v1_0.pdf;
34. Daugman.J., "How Iris Recognition Works", p.28, 2004, http://www.cs.columbia.edu/~belhumeur/courses/biometrics/2010/howirisrecognitionworks.pdf;
35. S.Shanawaz Basha, N.Musrat Sultana, "Spectral Minutiae Fingerprint Recognition System", http://www.idc-online.com/technical_references/pdfs/data_communications/Spectral%20Minutiae%20Fingerprint.pdf;
36. SHA-2, http://en.wikipedia.org/wiki/SHA-2;
37. Stevens.M., "Cryptanalysis of MD5 & SHA-1", 2012;
38. Simon., L., Silverman M, "A Practical Guide to Biometric Security Technology". http://www.computer.org/itpro/homepage/Jan_Feb/security3.htm;
39. Tsoklinova, M., The changes in consumption induced by the use of information and communication technologies, "Application of information and communication technology and statistics in economy and education" (ICAICTSEE-2012), University of National and World Economy, Sofia, 2013, pp. 518-525.

A Practical Approach to Testing the Execution Times of the Queries in Microsoft SQL Server

Mitko Radoev

Department of Information Technologies and Communications, Faculty of Applied Informatics and Statistics, University of World and National Economy

Abstract. The response time of an information system is one of its most important characteristics. One part of the response time is time for requests to the database. Knowing this time in advance is vital before implementing of the system in exploitation.

However, an accurate measurement of execution time of the queries is faced with many difficulties.

This study aims to propose a practical approach to objectively measure the execution time of database queries in an environment of Microsoft SQL Server, as part of the testing of information system before its exploitation.

Keywords: response time, query execution time, performance testing

1 Introduction

The response time of an information system is one of its most important characteristics. Sometimes this time is crucial for the normal exploitation of the system.

One part of the response time is time for requests to the database. Although most studies show that the execution time of the queries occupies only a small part of the overall response time [1, 2], it is important to know how much this time is. And even more important is to know this in advance, before the implementation of the system in exploitation, in order to eliminate potential problems.

The subject of this study is to clarify:

- What part is the execution time of database queries from the response time of the system?
- What should be included in the execution time of the queries, i.e. what should be measured?
- What are the conditions for the most objective measurement?
- How exactly to perform the measurements and which are the most appropriate tools in the environment of MS SQL Server?

1.1 Definitions

Response time: The elapsed time between the end of an inquiry or demand on a computer system and the beginning of a response; for example, the length of the time between an indication of the end of an inquiry and the display of the first character of the response at a user terminal. [3]

Transaction time: The total amount of time spent at network, client, and server to complete a transaction. [4]

Frontend time: Network and client-side time

Backend time: The time it takes the server to get the first byte back to the client [1]

Process time: The time required for the server to respond to a client request. [4]

Wait time: The time when the query waits for the CPU or other resources on the server to be processed.

Parse and compile time: The time required by the server for compiling, analyzing and creating execution plan.

Execution time: The time required by the server to execute the query.

1.2 Goals

The main goal of this study is to propose a practical approach to objectively measure the execution time of database queries in an environment of Microsoft SQL Server, as part of the testing of information system before its implementation.

Steps to achieve this goal:

- Determine the place of the execution time of the queries in the response time of the system;
- Determine what is included in the execution time of the queries;
- Define the conditions for maximum objective measurement;
- Identify appropriate tools in an environment of MS SQL Server.

2 Response time and query execution time

2.1 Components of the response time

The response time is composed of two main components:

- Frontend (client and network) time: The time it takes for the application to form a query to the server, the time required by the operating system and the network to transport the query and vice versa - the time it takes the operating system and the network to transport the result of the query and the time required to application to display the result;
- Backend (server-side) time: The time required by the server to prepare the query, to process it and to return the result.

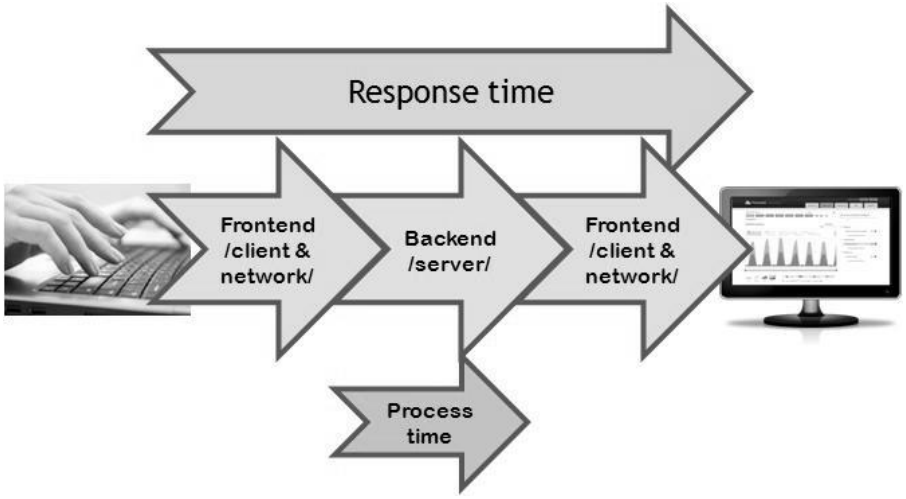


Fig. 13. Components of the response time

2.2 Components of the process time

The process time is composed of three main components:

- Wait time: The time during which the query is waiting to be processed by the server. There are two categories of requests - *runnable* (waiting for CPU) and *suspended*. Requests with suspended statuses are placed in a waiting list until the requested resources become available;
- Parse and compile time. Parse time is the time spent on checking SQL statement for syntax errors, dividing the command of components, and producing an execution tree. Compile time is time spent during compiling an execution plan from the execution tree.
- Execution time: Total time spent during execution of the query.

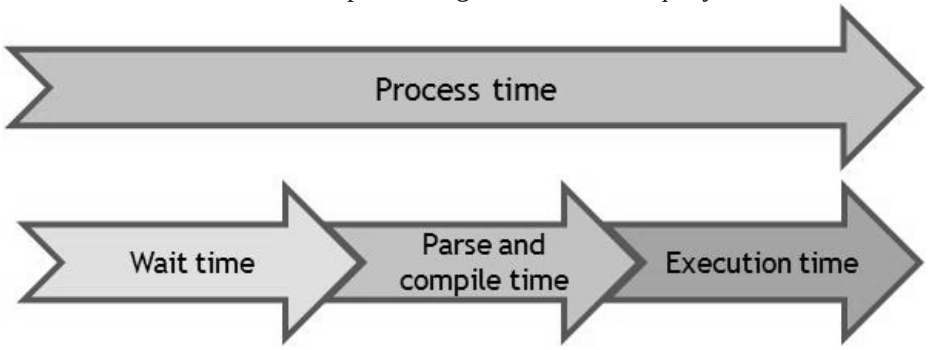


Fig. 14. Components of the process time

2.3 Query execution time

There are two indicators of execution time - *CPU time* and *Elapsed time*.

- CPU time is time used by CPU to complete a task. CPU time is just the time for which the CPU is busy executing the task;
- Elapsed time is the total time used by a task from start to its end. In most cases elapsed time is larger than CPU time, because it also includes time spent during I/O operations required by query. Elapsed time can be different for each execution because it depends upon other resources availability.

Later in the study under the *query execution time* will be understood only the *CPU time*.

3 Conditions and tools for measuring the execution times of the queries

3.1 Conditions for maximum objective measurement of the queries execution time

There are incidental factors influencing the execution time of applications, such as:

- Running background processes
- Server overload
- Query optimization and building different execution plans
- Resource constraints etc.

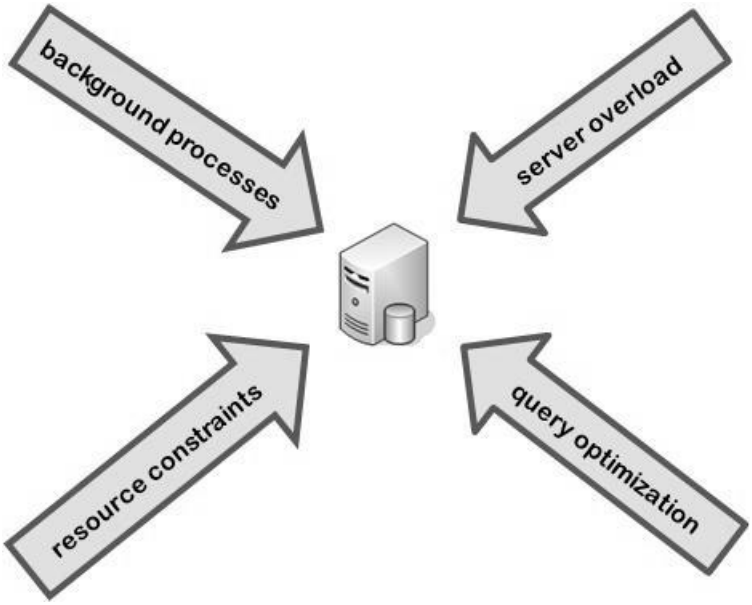


Fig. 15. Incidental factors

Because it is not possible to eliminate all the effect of the incidental factors, they should be limited to a minimum. Our objective in this study will be:

- Limitation of incidental factors
- Repeated measurements of the execution time of each query
- Statistical treatment of the results of the measurements;

3.2 Available tools in MS SQL Server

In MS SQL Server there are execution related dynamic management views and functions, such as:

- sys.dm_exec_query_stats
- sys.dm_exec_sql_text
- sys.dm_exec_query_plan

The view sys.dm_exec_query_stats returns aggregate performance statistics for cached query plans. It contains one row per query statement within the cached plan. Most important for the purposes of the study are the columns:

- execution_count - number of times that the plan has been executed
- total_worker_time - total amount of CPU time, reported in microseconds, that was consumed by executions of this plan
- total_elapsed_time - total elapsed time, reported in microseconds, for completed executions of this plan

4 An approach to testing the queries' execution time in Microsoft SQL Server

To ensure maximum objective measurement of the queries' execution time, we offer an approach to identify queries, execute them and process the results.

4.1 How to create and identify queries

First of all, a test database must be created with parameters close to the one expected in the actual exploitation of the system.

Then, queries corresponding to those in the real information system need to be created. Despite various specific parameters, these standard queries must be identified equally. One possible approach is, at the beginning of each query, to place a comment that identifies it, such as:

```
--Query #001
SELECT …
```

4.2 How to execute queries

Queries should be performed repeatedly with different parameters and in different order, as far as the order of execution has a significant effect on their performance. [5] The best option for this is to create an application that can run queries.

Before beginning the tests would be good to clear the statistics relating to the execution of queries, using database console command:

```
DBCC FREEPROCCACHE
```

4.3 How to process the results

After the tests necessary information about the execution times of the queries can be derived from the system views and functions **sys.dm_exec_query_stat**s, **sys.dm_exec_sql_text** and **sys.dm_exec_query_plan**.

For each query can be calculated the average execution time, for example:

```
SELECT 'Query #001' AS Query,
       SUM(s.execution_count)AS Count,
       SUM(s.total_worker_time) as Total_time,
       SUM(s.total_worker_time)/SUM(s.execution_count)
       as Average_time
FROM   sys.dm_exec_query_stats AS s CROSS APPLY
       sys.dm_exec_sql_text(s.sql_handle) AS t
WHERE  t.text like '--Query #001%'
```

References

1. Souders, Steve. the Performance Golden Rule. stevesouders.com. [Online] 2 10, 2012. http://www.stevesouders.com/blog/2012/02/10/the-performance-golden-rule/.
2. Podelko, Alex. Performance vs. Scalability. alexanderpodelko.com. [Online] 4 5, 2013. http://alexanderpodelko.com/blog/2013/04/05/performance-vs-scalability/.
3. ISO/IEC. ISO/IEC 2382-7:2000 Information technology -- Vocabulary -- Part 7: Computer programming. s.l.: International Organization for Standardization, Geneva, Switzerland, 2000.
4. Nguyen, Hung Quoc. Testing Applications on the Web: Test Planning for Internet-Based Systems. s.l.: John Wiley & Sons, 2001. ISBN 0-417-43764-6.
5. Modeling and Exploiting Query Interactions in Database. Mumtaz Ahmad, Ashraf Aboulnaga, Shivnath Babu, Kamesh Munagala. New York: ACM, 2008. CIKM '08 Proceedings of the 17th ACM conference on Information and knowledge management. pp. 183-192. ISBN 978-1-59593-991-3.

Common Architecture Design of a Business Information System for Performance Management of the Business Applications

Veska Mihova [1]

[1] UNWE, 1700 Sofia, Bulgaria
vmihova@unwe.bg

Abstract. With the development of information systems and the growth of managed and stored data arises a need for monitoring and management systems. The companies are increasingly interested in the performance of the systems which manage their business.This report is a result of current and previous research for systems' performance and more specifically the focus is on the database productivity in a system. The database is considered as a part of the entire business intelligent system. This report proposes common architecture of a business information system for performance management of a database. A series of steps, which the performance management system provides, will be described in the paper. We will also propose and describe tools which can implement these steps.

Keywords. Database monitoring, database administration, database performance, performance forecast.

1. Introduction

The main problems that business applications have are relevant to their productivity and performance.

Previous researches present an entire concept for management of database performance. The idea covers tasks and concepts that are very similar to those of business intelligent systems - data collection, storage history, producing new values based on the accumulated history, forecast. The previous reports take the idea of the architecture of business intelligent systems and develop it further to serve the idea of future forecasted environment, monitoring in forecasted state of the database.

The current report will propose a common architecture of a business information system for management of database performance based on the entire concept presented in previous reports.

The main purpose of this paper is to describe a sequence of certain functionality, which the system for database performance management would provide. Along with that, there are suggested and described some tools, which can help with the steps' implementation.

2. Common architecture

The prototype of the system consists of four component systems. They exist together under a single web interface. Of course, the implementation of each component system can be different, but for the presented prototype the author has chosen to develop a web interface using ASP.NET and C#.

The specific steps for execution, which the system for database performance management suggests can be classified by the component system, by which they are supposed to be performed.

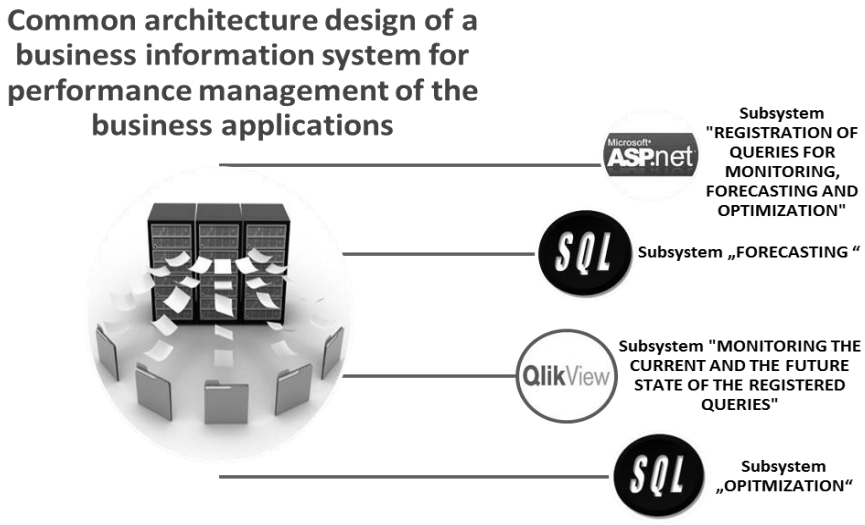


Fig. 16. Common architecture design of a business information system for performance management of the business applications

1. Subsystem "Registration of queries for monitoring, forecasting and optimization“

This part of the system is designed to work with the common properties of the database that will be monitored. This component system should also allow the user – database admin to specify business problems, represented in the system using as large queries, which the organization needs to optimize and monitor. The user should also be able to set critical thresholds for execution time of the queries.

The so-called “execution steps” briefly describe the main functionality of the component system “Registration of queries”:

- Creating stored procedures for generating storage structures for statistical data about the work of the database depending on different servers.
  Analysis of tables and views, which hold or access statistical data for the different database management systems respectively. Creating stored procedures corresponding to a selected server or database management system.

- Collection and management of data for generating a specific data warehouse.

This step is designed to include collecting of data necessary for creating a specific data warehouse in a prepared in advance universal storage structure.

The data warehouse, which will be generated, is defined as specific, because it is created according to some specific properties such as the database server, which is used for the database and some others.

Each database management system stores statistical data for the work of each database in various structures. These structures are usually tables that differ not based on the databases, but based on the server and the database management system, on which they are stored and used.

- Generating a specific data warehouse for storing statistics for the operation of the database, depending on the specifics of a database server;
  On this step the data warehouse is generated according to the data for the database, collected during the first two steps.

- Registration of server, database, method of authentication;
  This step includes collecting information for the name of the server of the database which will be monitored as well as the name of the database. Another important thing is to choose a way of identification for the server– windows or sql user.

- Registration of business processes for monitoring;
  The suggested method equalizes the registration of a business problem with registration of a critical for the business in terms of query execution time.

- Registration of sources which depend on the business processes;
  For research purposes there should be registered different data sources, which participate in the studied business process, in order to monitor, forecast and detect potential problems with the database performance depending on the given business process.

- Registration of performance values.
  The database administrator is responsible for determining critical thresholds for the execution time of a given query and in this way the process will be monitored according to its optimal values.

2. Subsystem „Forecasting“

This component system is the heart of the system, it is developed using the DBMS SQLServer and T-SQL. The user interface for using the implemented functionality is part of the common ASP.Net application.

This part of the system can be developed as well using diverse tool and technologies, for example – Data mining, Statistics packages, MS Excel, stored procedures, etc. For the current research the author has used stored procedures and the help of statistical methods for forecasting, detecting dependencies and calculating the correlation coefficients.

- Collection of statistics for execution time of queries;
- Generate forecast values for execution time of queries;
- Collection of statistics for database objects;
- Generate forecast values for the database objects;
- Calculation of correlation coefficients based on statistics;
- Generate forecast values for correlation coefficients.

- Periodic collection and generation of statistics and their forecast values, calculating the correlation coefficients and generating forecast values for correlation coefficients.

3. Subsystem "Monitoring the current and the future state of the registered business processes"

This part of the system allows the database administrator to monitor the system and its indicators through a dynamic dashboard. In this way the administrator will have a view containing all registered business processes, their current and forecasted state, the volumes of the data sources participating in these processes, as well as the execution time for each registered query in separate moments. All these current and forecasted values can be compared to critical thresholds, set for each of the business processes. What is more, in this dynamic dashboards can be monitored and detected dependences between the different data sources participating in a given business process. Monitoring of the calculated forecasted values for detecting correlations between the different sources is a feature of this subsystem too. In this way an accurate and timely decision can be made for detecting and preventing problems with the database performance.

Similarly to the other subsystems, this part can be developed in multiple ways, in the prototype it's chosen to be developed using Qlikview tool.

The following steps cover the main functionality of this component:

- Monitoring the execution time of queries registered for a specified period;
- Monitoring forecast values for execution time of queries registered for a specified period;
- Monitoring the values of database objects associated with the registered queries;
- Monitoring forecast values of database objects associated with the registered queries;
- Monitoring of the calculated correlation coefficients based on statistical data;
- Monitoring forecast values for correlation coefficients based on statistical data.

4. Subsystem „OPITMIZATION"

As all systems for monitoring and optimization of database performance follow the idea of analyzing diverse statistics about the objects in the database, this subsystem is aimed to organize the process of entering all the statistics collected by the rest of the subsystems into an empty virtual database. This means that this database will contain no data, just the statistics, in order to allow the query optimizers to propose accurate solution for optimization of the already forecasted values for volumes of data in the different objects in the database. In this way the administrators will be able to look one step ahead in the development of the database volumes, which they are responsible for.

- Load forecast statistics;
- Generating Execution plan based on forecast statistics.

3. Conclusion

In summary, in this paper a common architecture of a business information system for performance management of the business applications is presented. In the beginning the author presents the idea of a system and its subsystems. The author proposes architecture of

business information system in help of database administrators. The proposed solution gives the database administrator enough time for reaction to an occurring problem with the application's performance and respectively the database's performance.

Fig. 2. Better database's performance

As for future work, the author intends to present the tool using the concept of this paper.

References

1. Lane, Adrian. Database activity monitoring keeps watch over your data. http://searchsecurity.techtarget.com. [Онлайн] http://searchsecurity.techtarget.com/article/Database-activity-monitoring-keeps-watch-over-your-data.
2. Quarles van Ufford, Deborah. Business Intelligence: The Umbrella Term, VU Amsterdam. [Онлайн] 2002 r. http://www.few.vu.nl/stagebureau/.
3. Yuhanna, Noel, Penn, Jonathan и Smillie, Katie . Forrester Research:Research: The Forrester Wave™: Enterprise Database Auditing And Real-Time Protection, Q4 2007. [Онлайн] October 2007 r. http://www.forrester.com/The+Forrester+Wave+Enterprise+Database+Auditing+And+RealTime+Protection+Q4+2007/fulltext/-/E-RES41970?objectid=RES41970.
4. Mihova, V., Monitoring the future state of a database, UNITE 2nd Doctoral Symposium R&D in Future Internet and Enterprise Interoperability, 11-12.10.2012, Sofia.
5. Murdjeva A., Mihova V., Metadata for generating a specific data warehouse, International Conference on Application of Information and Communication Technology and Statistics in Economy and Education ICAICTSEE 2012 UNWE, Sofia , Bulgaria, 5-6 October 2012

Object-Oriented Architecture for Simulation of Complex Interdependent Systems Based on HLA Standard

Georgi Kirov[1], Valentin Stoyanov[1]

[1] ISER - BAS, Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria
g_tk@abv.bg

Abstract. This paper investigates an implementation of the Object-Oriented (OO) concepts into a High-Level Architecture (HLA)-networked simulation for transport systems, thus addressing sustainability of the HLA methodology into the future. The paper aims at developing an object-oriented layer (OOL) providing a high-level mechanism for HLA data exchange through local objects. It exploits the technical advantages provided by object-oriented programming, HLA object interface specification, and data-centric model while presenting an application for simulation of Air-Traffic Control (ATC), which is intended to serve as a testbed for ATC operators training.

Keywords. HLA/RTI simulation, object-oriented programming, integrated systems, transport systems.

1. Introduction

Over the last decades, the advent of new technologies has led to very complex and interdependent systems. They comprise a lot of basic facilities, services, information systems, and communication networks needed for the functioning of a community or society. The people in developed countries are becoming more and more dependent on these systems due to the critical operations and infrastructures they support, especially in the field of transport. One of the main prerequisites for the development of complex systems in transport is the tendency that the different main functions in society – such as transport management, air-traffic control, energy, communications, crisis management and health services – have developed a great deal of mutual interdependence. In addition to this, the *complex interdependencies* between the systems resulting in the mutual provision of services and use of common communications network prohibit the possibility to study these systems directly.

The modeling and analysis of interdependencies between complex systems in the transport is a relatively new and very important field of study. A number of simulation models have been developed and more are being developed for studying an individual aspect of the system components. In most cases, the simulation models are built based on the object-oriented approach. *Object-oriented simulation* has great intuitive appeal in applications because it is very easy to view the real world as being composed of objects [1]. Typical for the models is that integration among models does not exist. They do not easily address highly dynamic complex systems. *Distributed simulation technologies* are a paradigm to model dynamic, heterogeneous, and spatial distributed systems. They not only aim at speeding up simulations, but also serve as strategic technologies for linking simulation components of various types [2]. Although the contemporary distributed simulation technologies, and especially, HLA/RTI (High Level Architecture/Run Time Infrastructure)

standard has a standardized structure for object models, they do not completely correspond to common definitions of object models in OO techniques [3, 4].

The paper aims at developing an *object-oriented layer (OOL) for distributed simulation of interdependent systems in the transport providing a high-level mechanism for HLA data exchange through local objects*. By manipulating these objects (i.e., create, modify, delete), the applications can work directly at the HLA/RTI API. This will eliminate the complex network programming for HLA distributed simulations.

4. Problem statement

Given the current state of shrinking budgets and growing number of interdependent systems in the transport, it is obvious that there is a critical need for an architecture that exploits the technical advantages provided by object-oriented programming, HLA object interface specification, data-centric model, and layering (Fig. 1).

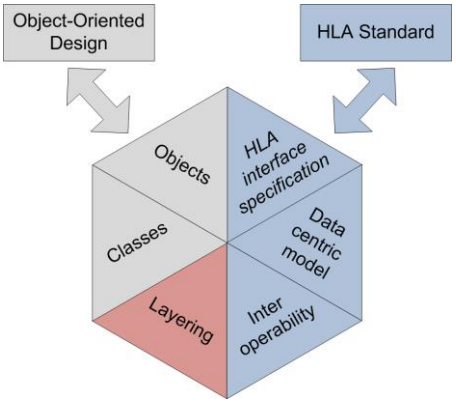


Fig. 1. Properties of the architecture for simulation of complex systems

1. Object-Oriented Simulation

Object-oriented programming (OOP) is a computer programming paradigm in which a software system is modeled as a set of objects that interact with each other. According to OOP methodology, the objects encapsulate data (attributes) and behavior (methods of behavior) [5]. In OOP, a Class definition is a user-defined type, which contains the members that enable class objects to have state and behavior. Data field members (member variables or instance variables) enable a class object to maintain a state. Other kinds of members, especially methods, enable a class object's behavior. Class instances are of the type of the associated class. OOP uses polymorphism. The purpose of the polymorphism is to implement a style of programming, in which objects of various types define a common interface of operations for users. This allows objects with the same interface to react differently depending on the nature or type of the object. For example, a method of a subclass can override a method of the base class with the same interface. Consequently, an instance of the subclass will differ from instance of base class in its behavior. Therefore, the specific operation can be handled differently depending on the type of the object on which it is invoked.

Object-oriented simulation is a design strategy where system designers think in terms of 'things' instead of operations or functions. It is concerned with developing an object-oriented model of the application domain. The identified objects reflect entities and

operations that are associated with the problem to be solved. The simulation system is made up of interacting objects that maintain their own local state (attributes) and provide operations (methods) on that state information. They hide information about the state and limit access to it. An OO design process involves designing the object classes and the relationships between these classes.

2. High Level Architecture (HLA) standard for distributed simulations

The High Level Architecture (HLA) [6] was developed by the U.S. Defense Modeling and Simulation Office (DMSO) as a successor to both Distributed Interactive Simulation (DIS) that supports real-time platform simulations, and Aggregate Level Simulation Protocol (ALSP), to provide simulation interoperability and reusability across all types of simulations [7]. HLA is the IEEE standard for a software architecture of interoperable distributed simulations. HLA aims to establish a common high-level simulation architecture to facilitate the interoperability of all types of models and simulations among themselves, as well as to facilitate the reuse of modeling and simulation (M&S) components. HLA allows experts to combine computer simulations into a larger simulation. For instance, the experts might want to combine simulations of complex systems in several different regions of the country. HLA can extend the simulation later by adding new models or simulations, for example new models of transport infrastructures.

The HLA defines a set of rules governing how different simulations (applications), now referred to as federates, interact with one another. The federates communicate via a communication environment called the Runtime Infrastructure (RTI) and use an Object Model Template (OMT), which describes the format of the data. A federation is a combined simulation system created from a set of federates that are interconnected with each other. Information exchange in the federation is based on a common object model, called Federation Object Model (FOM). It contains exchange data created by the federation developer that show the relationships between federates. It means, FOM defines object classes, their attributes and interaction classes that are commonly used and exchanged among the federates in the federation. The Simulation Object Model (SOM) is the model that defines objects, attributes and interactions in each federate that can be used from the other federates.

3. HLA object model

The *HLA object model* supports information exchange between federates within the federation. The exchange of information takes the form of objects and interactions. Federates communicate with their peers by sending interactions or updating object attributes. Federates do not communicate directly with each other and all communication is administrated by the RTI. A traditional HLA federate (Fig. 2) can be presented as an integrated program, consisting of a *simulation model* and *Local RTI Component (LRC)*. The simulation model is a physical, mathematical, or logical representation of processes and systems (user logic), whereas the LRC services it by interacting and synchronizing with other federates. Therefore, the simulation model performs local computing, while the LRC realizes information exchange for the model. It is very difficult to combine the above components in a normal federate, due to the tight coupling of the simulation model and the LRC [2].

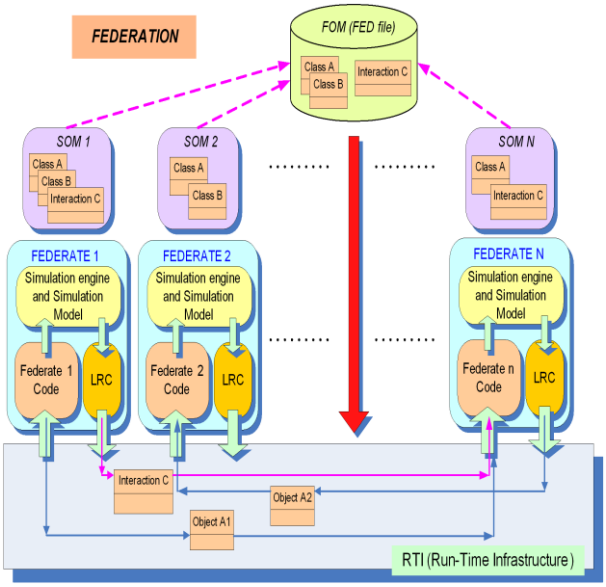


Fig. 2. Traditional HLA simulation

HLA object classes are comprised of attributes. Object classes describe types of things that can persist. Each object in a moment of time is characterized by a state, which is defined by a set of current values of its attributes. Federate, which manages an object (more precisely, the object attributes), may alter the state of the object by changing the attribute values. Through RTI services, the federate transmits the new values of the object to all federates in the simulation. In this case it is assumed that the federate updates the attributes. *Interactions classes* are comprised of parameters [8]. An interaction is a single action caused by a change in the state of an object from another federation. Interaction classes describe types of events. Objects are similar to the interactions in so much as the objects are comprised of attributes and the interactions are comprised of parameters. The basic difference between objects and interactions is persistence - objects persist, interactions do not.

In the HLA terminology, classes support information for a common description of the objects. The basic idea is to provide *data-centric organization*, which is an opportunity for information exchange between distributed applications. This concept differs significantly from the classical object-oriented methodology where the behavior is an integral part of the objects. Diversity of HLA and OOP object concepts consists in the fact that the HLA objects are defined entirely by the attributes, values of which are exchanged between federates. The responsibility for updating the attributes of an HLA object is distributed among different federates in a federation, whereas the OO objects encapsulate state locally and associate update responsibilities with methods that are integral part of the object's implementation [9].

In conclusion, the HLA FOM offers an object model that does not completely correspond to common definitions of object models in the object-oriented programming. The main purpose of the HLA FOM is to achieve interoperability between participants (federates) in the simulation, rather than between individual objects from different systems (federates). HLA doesn't support the transference of objects and their behaviors between joined federates. Therefore, there are fundamental differences between OO programming and HLA.

A number of assumptions about how a federate wants to use HLA services must be made in order to support these services in an OO API. On the other hand, it is also necessary to make a number of assumptions about the HLA interactions between federates in order to fully use object-oriented features such as method invocations [7]. The acquired experience in the HLA simulation technologies has called for carrying out in-depth analysis as to what extent the simulation mechanism is fitting the contemporary challenges and requirements. *The data-centric* nature of the HLA simulation is indicative of the following trends:

- The HLA technology makes it easy to connect applications, but not so easy to find, access, and work with the information from OO applications.
- Lack of an OO infrastructure of the distributed HLA simulation, providing effortless component integration.
- The distributed HLA simulation does not effectively address data abstraction and information hiding.

5. Object-Oriented Layer (OOL) for HLA simulation

Particular challenge to the simulation of complex systems in the transport comprises the rapid dynamic changes occurring in the operational environment, with multitude of variables and enormous amount of information to be processed and analyzed. In such circumstances all efforts must be focused on coordination, communication, information exchange and management at all levels. Therefore, the enhanced requirements call for applying a new approach, combining a methodological framework of the OOP with the capacities of the HLA standard for information exchange. The mentioned issues can find their appropriate solutions only after determination of various scientific, technical and technological requirements based on a comparative analysis of HLA and OOP.

The proposed OOL extends the HLA profile with an object-oriented view on a set of related HLA FOM object data thus providing typical OO-features such as navigation, inheritance and use of value types. The main goal of OOL is to provide functions and services for working with the traditional HLA API. Once this is done, we can write the business logic on-top of these abstractions. OOL can significantly simplify the implementation of HLA interfaces. Fig. 3 shows a high-level model of the OOL. Within the OOL, the rtiAmb contains customized libraries that accesses the standard RTI services and simplifies the design of the simulation model. Another member of the OOL, fedAmb, provides common callback mechanism to the programmer, and thus the RTI invokes functions from the OO user methods. The proposed OO approach aims at transferring objects between joined federates. It implies shared HLA objects to be presented as a local OO objects (C++), i.e. a HLA object instance to be presented as an object-oriented C++ object. For this purpose, the OO objects have to be mapped to the FOM data format. Therefore, HLA attributes have to be mapped to C++ class attributes.

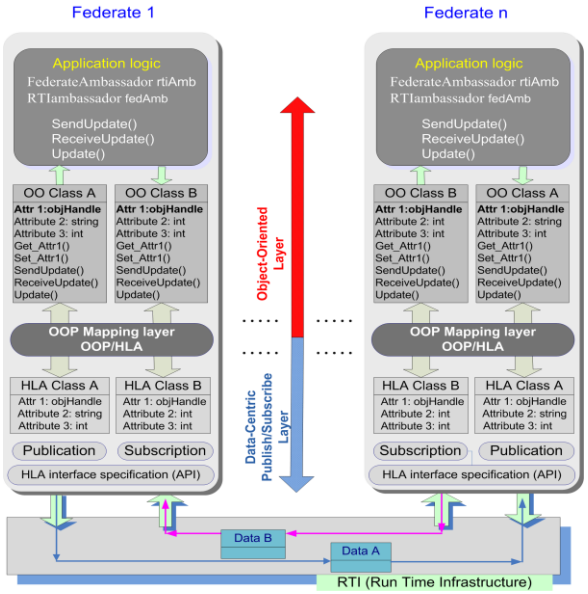


Fig. 3. Object-Oriented Layer for HLA simulation systems

OOL provides flexible methods to the user for packing update data, and leaves the transmission details transparent. The publishing federate uses an OO function sendUpdate() to send object data through the RTI. Each HLA object contains a member that presents a handle to the instance of the given type. It serves as a key, which is a unique representation of HLA data into the RTI middleware. The subscriber federate of the communication model receives published data for a given object. Then, the federate calls the function receiveUpdate() of the object distinguished by the key. The function maps HLA FOM object to an OO class, thus making it possible to reconstruct object-oriented views of the existing data models (mirror objects). Thus the mirror objects are automatically created at the receiving federates as object oriented instances.

The OOL implements a mechanism for an object transfer can be summarized in the following steps:

- Presentation of a HLA object instance as an object-oriented C++ object.
- Coding the C++ class attributes into a network format (bytes) by a serializing mechanism.
- Transmission of the serialized object attributes through the RTI by an OO function.
- Reconstruction of the object at the receiving federation according to the existing data models (mirror objects).

6. Reference OO Model For HLA Objects

For the creation of the OO HLA simulation model a DEVS (Discrete Event System Specification) formalism is used, because it enables the modular assembly and helps to reuse components [10].

1. DEVS formalism

The DEVS formalism was introduced in the late seventies by Bernard Zeigler as a basis for the compositional modeling and simulation of discrete-event systems [11]. System behavior in DEVS formalism is described using input and output events as well as states:

$$DEVS = (X, Y, S, t_a, \delta_{int}, \delta_{ext}, \lambda) \tag{1}$$

where: X is a set of input events, Y is a set output events, S is a set of sequential states, t_a is a time advance function, δ_{int} is the internal transition function, which defines how a state of the system changes internally (when the elapsed time reaches to the lifetime of the state), δ_{ext} is an external transition function which defines how a state of the system changes by external events, and λ is an output function.

$$t_a : S \to T^{\infty} \tag{2}$$

$$\delta_{\text{int}} : S \to S \tag{3}$$

$$\delta_{ext} : Q \times X \to S \tag{4}$$

$$Q = \{(s, t_e) \mid s \in S, t_e \in (T \cap [0, t_a(s)])\} \tag{5}$$

where Q is a set of total states, and t_e is the elapsed time since the last event. The coupled model of DEVS formalism provides a method of assembly of several atomic models to build complex systems [12].

The proposed OO HLA architecture is an extension of the HLA by including DEVS principles in interaction between federations. The design of the OO HLA simulation is inspired by the DEVS formalism, where a coupled model is composed of multiple models. Similarly, a HLA federation is analogous to a coupled model, while a federate (model) is to an atomic model. OO HLA model constructs the relationship between elements in the conceptual DEVS model and the data-centric elements of the HLA object definition:

- HLA published interaction and object classes are referred to in δint because they are used to produce output events from a model.
- HLA subscribed interaction and object classes are referred to in δext because they are used to inject input events into a model.
- The time advance function is identical to t_a because they both provide the time advance.

2. Object Oriented Model For HLA Objects

The object-oriented concept of the HLA architecture is based on a strict hierarchical organization of classes that satisfies the requirements for building DEVS systems based on the HLA standard. The reference model of the OO HLA object is a C++ class that sets a standard for building the HLA objects (Fig. 4). The Basic class StandAlone provides an abstract interface that is used to access each of the derived objects. It contains features that allow the HLA classes to be integrated into the simulation environment. StandAlone has virtual methods and data members that OO HLA objects must inherit to be able to participate in the information exchange. The virtual methods provide an implementation of the DEVS

model including the initialization, state maintenance, receiving events, outputting actions, and time advance and so on (Fig. 4):

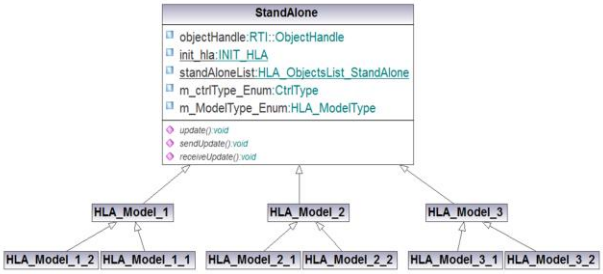


Fig. 4. Hierarchy of classes in the implementation of the HLA models

- receiveUpdate() – external transition function. If the attributes for the subscribed object classes are received before the time step is expired, it changes the state of the HLA model.
- update() – internal transition function. If no attributes are received, after the time step is expired, it changes the state of the HLA model.
- sendUpdate() – the function sends the attributes of the HLA model, calculated by the internal transition functions.

The main characteristics of the models (data and methods) are inherited from the class StandAlone. The application can manipulate these objects (i.e., create, modify, delete) using the object oriented language (C++). The OO HLA model provides mechanism to manage incoming updates and application-made modifications. Each OO HLA model is a C++ class that describes the behavior of an object. It ensures compliance with the requirements for building OOL which allows the exchange of objects, not just structures.

The state variables are recalculated on an each time step of the simulation time by the function update(). The updated values of the state variables are sent to the RTI environment by SendUpdate() where all subscribing applications can get them. When an HLA update is received the corresponding mirror object is updated, enabling the application to receive the value whenever needed.

7. Example for building a model of HLA model

To verify the effectiveness of the OOL, a case study was done on an integrated simulation of an Air Traffic Control (ATC) system. The scenario for this case study attempts to incorporate realism and flexibility. The purpose of the integrated simulation is to observe how the objects behave when unordinary events occur and potentially cascading effects, to test the effectiveness of the emergency response plans, for helping train response personnel, and for vulnerability analysis.

The integrated simulation system is created from a set of models that are interconnected with each other. The proposed simulation system consists of several federates. All simulated entities, such as aircrafts, radars, speech navigator, viewer, and analyzing tool are referred to as *simulation models* (Fig.5). They are developed following the specifications given into the reference model of OO HLA object.

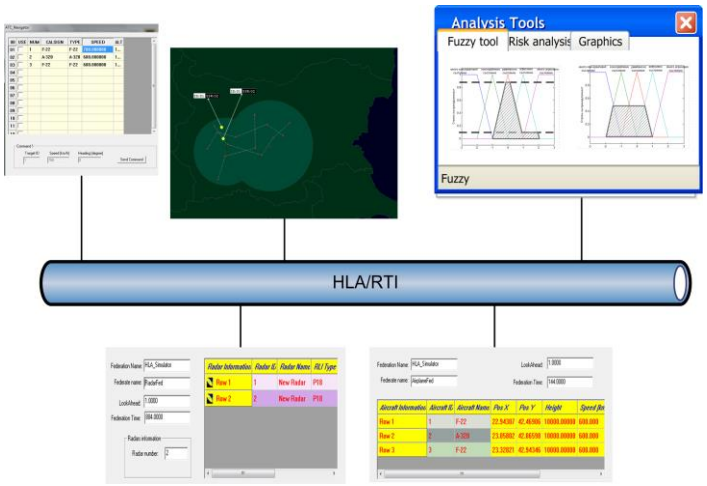


Fig. 5. HLA Air Traffic simulation

The simulation models consist of C++ code that accesses communication services provided by the RTI. This mechanism allows communication between simulation models based on HLA standard through RTI infrastructure. The communication between the simulated objects in the integrated simulation is based on a FOM. It contains exchange data created by the developers that show the relationships between the models.

The Aircraft federate simulates air traffic based on data provided by an user. In the example, all HLA Aircraft objects recalculate the new position of the Aircrafts by the function Update() in the each time step. The updated values of the state variables are sent to the HLA RTI environment by SendUpdate() where all subscribing applications can get them using the same interface and the same FOM.

The Radar federate simulates the operation of various radars. This is based on pre-built models for radars, which are integrated into the program. The Radar receives information about the position of an aircraft and determines if it is in the visibility zone, and then sends them through RTI to the other federates.

The viewer federate is developed to provide an integrated display environment. It can act as a passive recipient and display simulation data from the rest of the simulation system. The viewer uses databases to find geographic coordinates based on the given identifier of navaid or airport.

The analysis tool is an important part of the simulation system because it provides decision making procedures for solving traffic problems using different decision criteria and fuzzy multicriteria optimization. It describes the elements in the analysis of decision alternatives and choices, as well as the goals and objectives that guide decision-making. The speech navigator federate provides procedures for management of the aircrafts with voice commands. It supports the ATC operators training.

8. Conclusion

The paper presents an object-oriented layer for simulation of complex-interdependent transport systems based on an extended HLA integration architecture. The concept relies on the idea of providing a structural methodology for the use of the object-oriented method to design, develop, and realize an object model of a HLA simulation system, and obtain the interoperation and reuse of the simulation federations. It allows intensive information

exchange in a comprehensible format among distributed participants while following the principles of information security, regulated information access and coordination. The proposed OOL allows HLA simulation systems to be integrated easily into larger network-centric systems [13]. The commonly used subset of the full HLA functionality directly matches the object-oriented constructs of the OOL. The potential advantages of the proposed architecture are evident: increased flexibility, building on existing software and communications standards, and maximisation of the use of existing simulation models. The developed ATC simulation shows that the OOL reduces the lines of code that need to be written for a HLA application.

References

1. J. A. Joines and S. D. Roberts, ”An introduction to object-oriented simulation in c++”, Proceedings of the 1997 Winter Simulation Conference, S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, Eds., 1997, pp. 78-85.
2. D. Chena, S. J. Turner, W. Cai, and M. Xiong, “A decoupled federate architecture for high level architecture-based distributed simulation”, J. Parallel Distrib. Comput., vol. 68, 2008, pp. 1487–1503.
3. IEEE Standard for Modeling and Simulation (M&S), IEEE Std 1516.2-2000, 2010.
4. Tolk, “HLA-OMT versus Traditional Data and Object Modeling”, 2001 Command and Control Research and Technology Symposium, Annapolis, Maryland, June 2001.
5. H. M. Deitel and P. J. Deitel, “How to Program C++,” Prentice-Hall, Inc., ISBN 0-13-117334-0, 1994.
6. W. Cai, Z. Yuan, M. Y. Low, and S. J. Turner, “Federate migration in HLA-based simulation”, Future Generation Computer Systems, vol 21, 2005, pp. 87–95.
7. B. Möller and F. Antelius, “Object-Oriented HLA - Does One Size Fit All?”, Proceedings of 2010 Spring Simulation Interoperability Workshop, 10S-SIW-058, Simulation Interoperability Standards Organization, April 2010.
8. F. Kuhl, R. Weatherly, and J. Dahmann, “Creating Computer Simulation Systems: An Introduction to the High Level Architecture”, Prentice Hall PTR, 1999, ISBN 0130225118.
9. IEEE, P 1516, “Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) – Framework and Rules”, February 2000.
10. M. M. Gutierrez and H. P. Leone, “DE2M: An environment for developing distributed and executable enterprise models”, Advances in Engineering Software, vol. 47, 2012, pp. 80–103.
11. J. H. Kim and T. G. Kim, “Hierarchical HLA: Mapping Hierarchical Model Structure into Hierarchical Federation” *M&S-MTSA'06*, Ottawa, Canada, July 2006, pp. 75 – 80.
12. B. Chen, R. Zhong and H. Vangheluwe, “Integrating Base Object Model Components into DEVS-based Simulation”, Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, vol 7(4), 2010, pp. 241–256.
13. C4ISR Architectures working group, Department of Defense, C4ISR Architecture Framework Version 2.0, 1997.

Soft Computing Model for Assessment of the Functional Effectiveness of the Enterprise

Yuliyan Velkov[1], Georgi Kirov[2]

[1] International Business School, Botevgrad, Bulgaria, e-mail: julianvelkov@yahoo.com

[2] Institute of Systems Engineering and Robotics, BAS, Bulgaria, e-mail: kirov@iser.bas.bg

Abstract. A soft computing model has been developed for evaluation of the enterprise's functional effectiveness. Based on the intelligent technologies and the information, generated in management accounting, it implements a systematic approach to the six indicators of the status and changes in assets and their sources. The model examines and analyses, through fuzzy sets and fuzzy logic, the traditional measures of operating assets and debt. It gives an expert interpretation of the financing structure. In terms of specific objectives, set out in the short-term management, comprehensive assessments the competitive business opportunities have been made. A complex fuzzy expert system for multi-criteria assessment has been developed for evaluation of the individual functional advantages of the enterprises. It has been created a model of the designed soft computing system, which has been simulated. The obtained results were examined, interpreted and analyzed.

Keywords. effectiveness, management, intelligent technologies, fuzzy sets and fuzzy logic, soft computing

1. Introduction

The enterprise's functional effectiveness gives a possibility for creation of utility on the basis of a specified value. In the evaluation of the functional effectiveness various measurements of the business activity and the financial position gain expression. Among the set of indicators forming the functional opportunities of the individual enterprise the key indicated ones are market share, rate of sales, turnover of resources and obligations, effectiveness of the costs, revenues, assets and equity [1, 3, 8]. The leading guidelines for the study of the operating activity and financial position are implementation (sales), management, investment and financing. Hence, on the basis of marketing, accounting and other past from historical perspective information knowledge with a focus to the short-term perspective is created. Very characteristic of the functional effectiveness is its strictly dynamic nature and its narrow interconnection with the tactical objectives of the management. Therefore, for the assessment of the functional effectiveness of the production plant it is necessary to develop a systematic approach that encompasses a large number of measures of the assets, liabilities and equity [2]. It is also necessary to have justification of the methodological tools, allowing for integration of the accounting information and the verbally defined hypotheses for the operation and financing of the enterprise.

The aim of this article is to create a soft computing model intended for complex assessment of the functional effectiveness of the enterprise. In is necessary to integrate accounting information for current assets and enterprise's financial position with knowledge

about business management, as well as to use intelligent methods of fuzzy sets and fuzzy logic.

2. Methods

The soft computing model for assessment of the functional effectiveness of the enterprise reproduces a complex expert system, designed and operating in an environment of uncertainty [5, 6]. It takes into account the lack of absolute accuracy (certainty) in the description of the objects and interfaces forming the functional opportunities. The conceptual bases of the soft computing model are the intelligent technologies - fuzzy sets and fuzzy logic [9]. Together with this a modern environment for improvement of the image of the operational functioning and capital structure is built up. Also there are conditions established for fusion between the classic, precise and straightforward financial measurements and expertise across the field of management, accounting, economic theory and analysis [7, 10].

The following requirements are raised towards the soft computing model for assessment of the functional effectiveness of the enterprise:

- To create a multi-criteria idea of the functional opportunities of the enterprise.
- To develop integrated knowledge for the active manifestation of the value invested in the operating activities in the enterprise.
- To build a complex knowledge of the passive manifestation of the value financing the operating activity in the enterprise.
- To improve the methods of business management by implementing smart technologies - fuzzy sets and fuzzy logic, in the traditional ways of management analysis.

The soft computing model for complex assessment of the enterprise's functional effectiveness is a mathematical image of the fuzzy expert system (Figure 1.). Six predetermined indicators are examined and three fuzzy indicators are introduced in it.

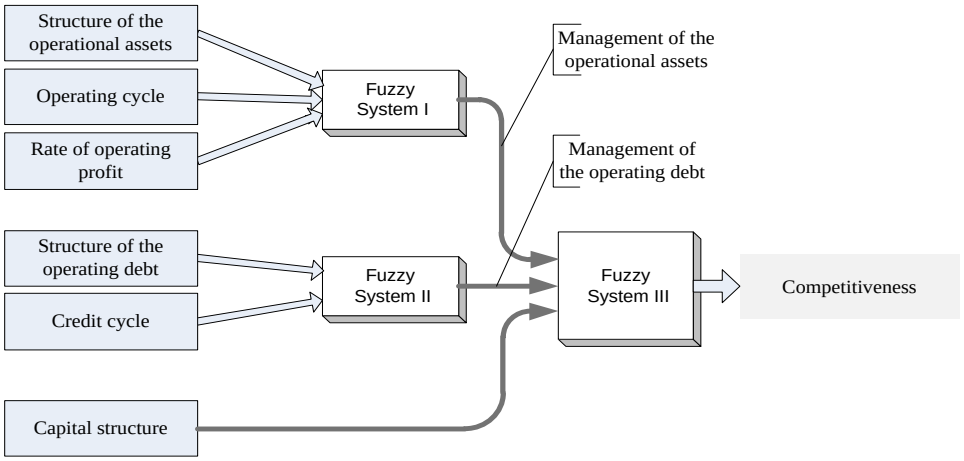


Figure 1. A schematic diagram of the Soft computing model for assessment of the functional effectiveness of the enterprise.

A. Determined indicators.

The determined indicators for assessment of the functional effectiveness are models of links with uniform follow-up between the preconditions and the outcome. They are calculated on the basis of the information, which is created mainly in the accountancy. The indicators of determined modelling are measured with an accurate value by means of a well-known formula. Their disadvantage is that they value individual sides of the operating activity, i.e. they give a limited account of the interrelationship between the characteristics of the operating assets, short-term liabilities and equity. In the proposed soft computing model the assessment of the functional effectiveness is formed on the basis of the following determined indicators [8]:

A-1. Structure of the operational assets. It expresses the ability for short-term operation. It focuses the study of the effectiveness on the current resources, intended for the production and marketing of production:

$$\frac{\text{Inventories} + \text{Receivables from Sales}}{\text{Operational Current Assets}} \tag{1}$$

A-2. Operating cycle. It creates a picture of the functioning of the enterprise and indicates the duration of the turnover realized by the operational short-term assets – inventories and receivables from sales:

$$\begin{pmatrix}\text{Duration of 1 turnover}\\ \text{of the Inventories}\end{pmatrix} + \begin{pmatrix}\text{Duration of 1 turnover}\\ \text{of the Receivables from Sales on Credit}\end{pmatrix} \tag{2}$$

Whereas:

$$\begin{pmatrix}\text{Duration of 1 turnover}\\ \text{of the Inventories}\end{pmatrix} = \frac{\overline{\text{Inventories}}}{\text{Material costs}} \times \text{Duration of the Reporting Period} \tag{3}$$

$$\begin{pmatrix}\text{Duration of 1 turnover}\\ \text{of the Receivable from}\\ \text{Sales on Credit}\end{pmatrix} = \frac{\overline{\text{Receivables from Sales on Credit}}}{\text{Sales on Credit}} \times \text{Duration of the Reporting Period} \tag{4}$$

A-3. Rate of operating profit. It evaluates the effectiveness of the ordinary activities of the enterprise. It shows the financial result (before interest and taxes), which is contained in a unit of revenues from sales of production, i.e. it measures the operating profit/loss separately from the influences that the available financial assets and liabilities have on it:

$$\frac{\text{Earring Before Interest and Taxes}}{\text{Revenues from Sales of Production}} \tag{5}$$

A-4. Structure of the operating debt. It characterizes the financing of the operating activity with funds payable to the suppliers:

$$\frac{\text{Obligations to suppliers}}{\text{Trade obligations}} \tag{6}$$

The commercial obligations are an amount of all debts, which are directly related to the operating activities - obligations for wages, social security and other. The amounts owed on obtained bank, bond and other loans are not an element of trade obligations i.e. the amounts resulting from assumed financial liabilities.

A-5. Credit cycle. It estimates the rate of repayment of the obligations on supplies. It shows the average time, which passes between the emergence and repayment of a single unit of commitment to the suppliers:

$$\frac{\text{Obligation to Suppliers}}{\text{Purchases on Credit}} \times \text{Duration of the Reporting Period} \tag{7}$$

A-6. Capital structure. It characterizes the origin (ownership) of the funds in the enterprise. It represents an assessment of the financial autonomy and expresses the degree, to which the creditors are secured:

$$\frac{\text{Equity Capital}}{\text{Borrowed Capital}} \tag{8}$$

B. Fuzzy indicators.

The fuzzy indicators for assessment of the functional effectiveness of the enterprise are introduced to reproduce human thinking [4, 9]. They are models of links, in which there is no determined dependency between the preconditions and the result. The content and the meaning of the fuzzy indicators used for the assessment of the effectiveness are reflection of expertly set theses. That is why the indicators appear to be an embodiment of the individual management estimates for the operating activities in the enterprise, for the policy on investment and financing, for the susceptibility to taking risk and others.

B-1. Management of the operational assets. It creates a multi-aspect idea of the resources in the ordinary activity. This indicator synthesizes the estimates of the Structure of the operational assets, the Operating cycle and the Rate of operating profit [8]. In the created soft computing model three hypotheses for management of the operational assets are defined: ***First hypothesis*:** Aggressive management of the operational assets. This management is judged as *highly profitable and very risky*. ***Second hypothesis:*** Moderate management of the operational assets. This management is defined as *moderately profitable and of average risk*. ***Third hypothesis:*** Conservative management of the operational assets. This management qualifies as *low profitable and of little risk*.

B-2. Management of the operating debt. It offers knowledge for financing of the operating activity by summarizing the estimates of the Structure of the operating debt and the Credit cycle [8]. In the proposed soft computing model three hypotheses for the management of the operating debt are developed: ***First hypothesis:*** Aggressive management of the operating debt. This management is judged as *highly profitable and very risky*. ***Second hypothesis:*** Moderate management of the operating debt. This management is defined as

moderately profitable and of average risk. ***Third hypothesis:*** Conservative management of the operating debt. This management qualifies as *low profitable and of little risk*.

B-3. Functional effectiveness. It creates a multi-criteria idea of the functional opportunities of the enterprise. It forms knowledge of the effectiveness and security of business operation on the basis of six predetermined indicators - Structure of operational assets, Operating cycle, Rate of operating profit, Structure of operating debt, Credit cycle and Capital structure. In the proposed soft computing model three hypotheses for the competitiveness are developed: ***First hypothesis:*** Strong functional effectiveness. This management is judged as *highly profitable and very risky*. ***Second hypothesis:*** Average functional effectiveness. This management is defined as *moderately profitable and of average risk.* ***Third hypothesis:*** Low functional effectiveness. This management qualifies *as low profitable and of little risk*.

3. Design of the Soft computing model

The design is implemented under the standard algorithm imposed in the building up of complex expert systems [4, 5]. It is carried out by successive construction of three fuzzy systems (FS):

FS 1: Comprehensive assessment of the operational assets of the enterprise.

Stage 1: Fuzzing of the input variables of FS 1.

It represents transformation of the exact values measuring each of the three input variables – Structure of operational assets, Operating cycle and Rate of operating profit, in fuzzy estimates. It defines the output variable of Management of the operational assets. For the purposes of fuzzing four fuzzy linguistic variables are developed (Figure 2).

Stage 2: Calculation of the rules of FS 1.

It is realized by the definition of logical rules, by which the term-multitude of the three fuzzy variables – inputs of the Soft computing model are associated with the term-multitude of the output variable. In the calculation of the rules the fundamental understanding is the following one: *With the optimum structure of operational assets, an Operational cycle of normal length and normal Rate of operating profit, the Management of the operational assets is of moderate risk and moderately profitable*. This Management is described as Equilibrium.

Any deviation from the normal estimates of the input indicators is accompanied by a corresponding change in the yield and risk, i.e. with the corresponding new qualifications of the Management of the operational assets.

Stage 3: Defuzzing of the outcome of FS 1.

The defuzzing of the result is carried out according to the method Geometrical centre of gravity:

$$COG = \frac{\sum_{i=1}^{n} \mu(x_i) \times xi}{\sum_{i=1}^{n} \mu(xi)} \qquad (9)$$

Whereas:

x_i - Element of the multitude-medium of the linguistic variable Management of the operational assets; $\mu(x_i)$ - Degree of affiliation of each element x_i

Hence, the accurate, unambiguous assessment of the Management of the operational assets is reached. FS 1 is modelled and simulated in an environment of Matlab by using the simulation tools Fuzzy Toolbox and Simulink. Figure 3. shows the dependency of the Management of the operational assets from the Structure of the operational assets, the Operational cycle and the Rate of operating profit.

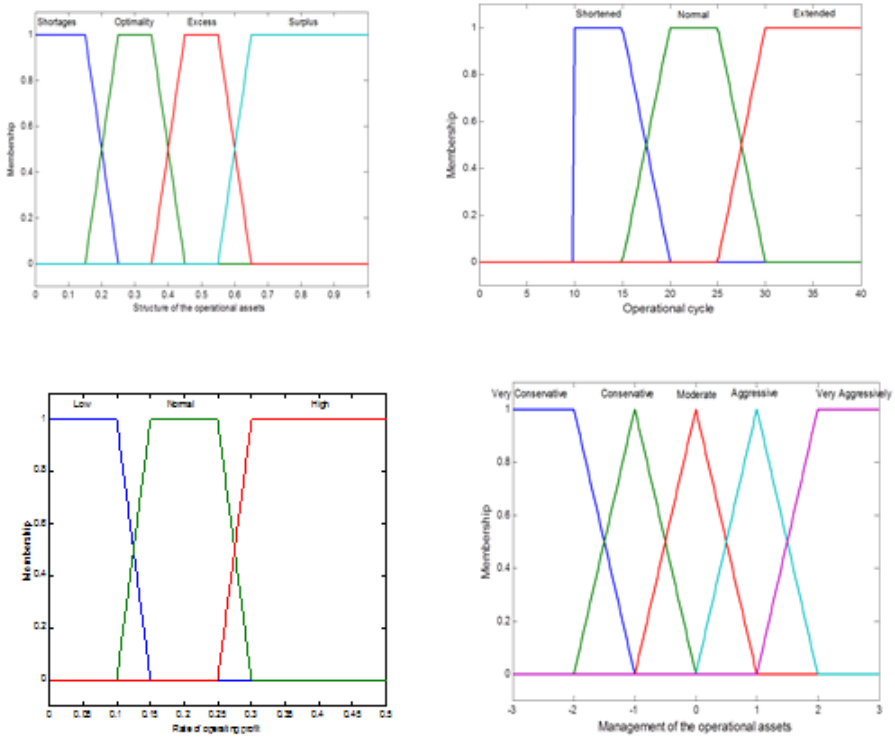


Figure 2. The linguistic variables Structure of the operational assets, Operating cycle, Rate of operating profit, Management of the operational assets.

FS 2: Comprehensive assessment of the operating debt of the enterprise.

The design includes the three stages: ***Stage 1: Fuzzing of input data.*** For the purposes of fuzzing three fuzzy linguistic variables are developed: Structure of the operating debt, Credit cycle and Management of the operating debt. ***Stage 2: Calculation of the rules.*** In the calculation of the rules the fundamental understanding is the following one: *With the optimum structure of operational assets, an Operational cycle of normal length and normal Rate of operating profit, the Management of the operational assets is of moderate risk and moderately profitable*. ***Stage 3: Defuzzing of the result***. It is carried out according to the method Geometrical centre of gravity.

FS 2 is modelled and simulated in an environment of Matlab by using the simulation tools Fuzzy Toolbox and Simulink. Figure 4 shows the dependency of the Management of the operating debt from the Structure of the operating debt and the Credit cycle.

FS 3: Comprehensive assessment of the competitiveness of the enterprise.

FS 3 is of the second level. It synthesizes the estimates of the Management of the operational assets and the Management of the operating debt, which are obtained at the outcome respectively to the FS 1 and FS 2 and integrates them with the assessment of the capital structure.

The design includes the three stages: ***Stage 1: Fuzzing of input data.*** For the purposes of fuzzing four fuzzy linguistic variables are developed: Management of the operational assets and Management of the operating debt (they are already done in the design respectively of FS 1 and FS 2), Capital structure and Management of the functional effectiveness. ***Stage 2: Calculation of the rules.*** It is implemented in the same way as the calculation of the rules of FS 1 and FS 2. In the definition of the rules the fundamental assumption is that in *moderate Management of operational assets, moderate Management of the operating debt and medium Capital structure, the Competitiveness of the enterprise qualifies as an Average*. ***Stage 3: Defuzzing of the result***. It is performed in the same way as the defuzzing of the result, which is already implemented in the design of FS 1 and FS 2 – according to the method Geometrical centre of gravity.

FS 3 is modelled and simulated in an environment of Matlab by using the simulation tools Fuzzy Toolbox and Simulink. Figure 3, Figure 4, and Figure 5 show the dependency of the Management of the functional effectiveness from the Management of the operational assets, Management of the operating debt and Capital structure.

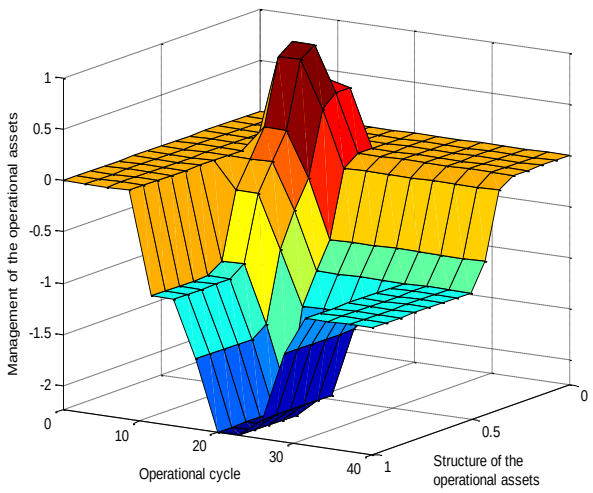


Figure 3. The dependency of the Management of the operational assets from the Structure of the operational assets, the Operational cycle and the Rate of operating profit.

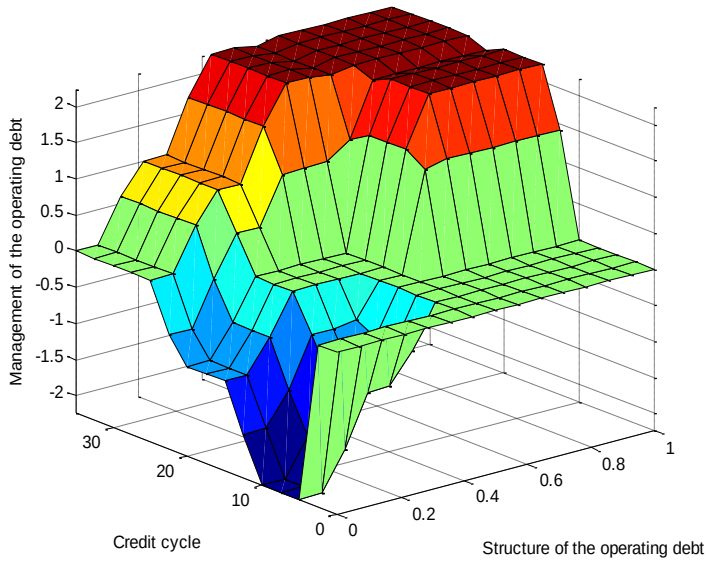


Figure 4. The dependency of the Management of the operating debt from the Structure of the operating debt and the Credit cycle.

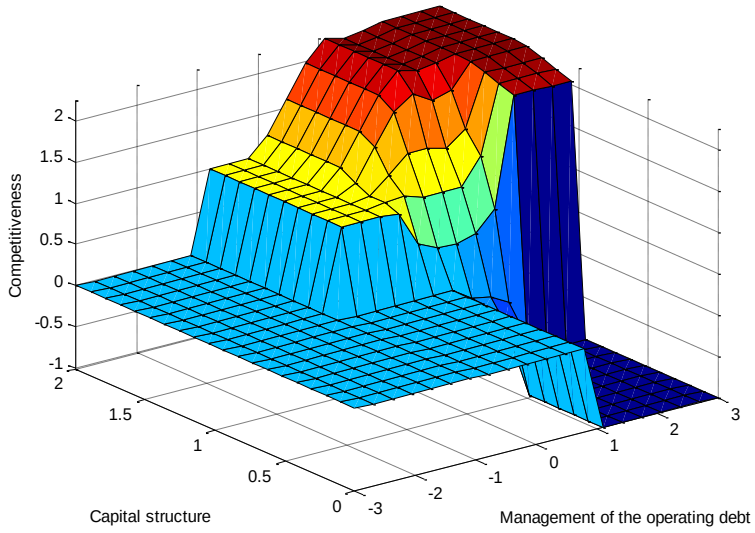


Figure 5. The dependency of the Management of the functional effectiveness from the Management of the operational assets, Management of the operating debt and Capital structure.

4. Implementation and testing of the Soft computing model

The soft computing model, intended for assessment of the enterprise's functional effectiveness, is realized and tested by the development of a testing example. In it the effects of the six scenarios on the Functional effectiveness of enterprise "A" are estimated (Table 1.).

Table 1. Scenarios on the Functional effectiveness of enterprise "A"

	SOA	OC	ROP	SOD	CC	CS	MOA	MOD	FE
Really position	0.11	27	0.13	0.12	11	1.05	-0.99	-1.72	1.50
Scenario 1	0.32	27	0.13	0.12	11	1.05	-0.99	-1.72	1.50
Scenario 2	0.32	17	0.13	0.12	11	1.05	0.12	-1.71	0.41
Scenario 3	0.32	17	0.27	0.12	11	1.05	1.10	-1.71	0.66
Scenario 4	0.32	17	0.27	0.37	11	1.05	1.10	-0.46	0.76
Scenario 5	0.32	17	0.27	0.37	24	1.05	1.10	1.6	1.46
Scenario 6	0.32	17	0.27	0.37	24	1.45	1.10	1.6	1.90

Whereas: SOA – Structure of the operational assets; OC – Operating cycle (days); ROP – Rate of operating profit; SOD – Structure of the operating debt; CC – Credit cycle (days); CS – Capital structure; MOA – Management of the operational assets; MOD – Management of the operating debt; FE – Functional effectiveness

Conclusions.

The Soft computing model of assessment of the functional effectiveness of the enterprise offers significant opportunities for the development of the modern economic management:

- It improves the consistency of the study of the enterprise's operating activities that together with the accounting data presented in the management reports uses qualitatively set information - for the macroeconomic environment, markets - of raw materials and product, the trading practice, etc.;
- It enhances the complexity of knowledge for the effectiveness of the enterprise - on the basis of expertly formulated rules, several measurements of the operational resources and operating debts, as well as the indicator for the capital structure, are encompassed in a single, summary evaluation.
- It expands the possibility for the analysis of the factor influences on the effectiveness of the enterprise.
- It creates conditions for the approximation between the accounting estimates, which show different aspects of the operational investment and financing, as well as the individual approach for the short-term management of the business.

- It offers opportunity for the use of the evaluation of functional effectiveness for macroeconomic purposes - for industrial, sectorial and other comparisons, for study of the effects, which certain economic decisions would have on the effectiveness of enterprises, etc.

The built soft computing model is modularly extensible. It allows change of the investigated characteristics of the investment and financing, on the basis of which a perception for the effectiveness of the enterprise is formed; and there is a possibility to set the new logical links between the indicators of assets, debt and equity and the assessment of the operating activity. The design and implementation of the proposed model are subject to further development and improvement.

References

1. Adamov, V. Finance of the company, "ABAGAR", Veliko Tarnovo, 2006
2. Branham, M. "Measuring the Components of Competitiveness", Competitive Benchmarking Association, Michael Branham&Cowles Business Media, 1997
3. Danailov, D, Finance management of the company, "Lyuren", Sofia, 1994.
4. Negoita,C., D. Ralescu. Applications of Fuzzy Sets Systems Analysis, Birkhauser Verlag, Basel, 1975
5. Sankar, K., M. Sushmita. Fuzzy Logic and Expert Systems Applications, 1998.
6. Sgurev, V., R. Pavlov. Expert systems, "Nauka I izkustvo", Sofia, 1989.
7. Stancheva, A. et all. Competitiveness behaviour of the company, "Nauka I izkustvo", Varna, 2006.
8. Trifonov, T., S. Trifonova. Finance analysis of the company, II part, "Siela", Sofia, 2001.
9. Vasileva, M. Fuzzy sets. Theory and practise, Shumen, 2008.
10. Velev, M. Assessment and analysis of the competitiveness of the company, „Softtreid", Sofia, 2004.

Social Network Analysis as a Tool for e-Government Maturity Measurement

Mimoza Bogdanoska-Jovanovska[1], Violeta Manevska[2]

[1] St. Kliment Ohridski University, Faculty of Administration and Information Systems Management, Partizanska bb, MK-7000 Bitola, Macedonia
mimozajovanovska@yahoo.com

[2] St. Kliment Ohridski University, Faculty of Administration and Information Systems Management, Partizanska bb, MK-7000 Bitola, Macedonia
violeta.manevska@gmail.com

Abstract. E-government benchmarking becomes interest topic in the last fifteen years. Researchers and practitioners have established a number of indicators integrated in indexes used in the process of e-government benchmarking for comparative and estimation of e-government development. Many of the indicators have been introduces with a narrow focus, poorly designed, by using simple statistical operations in the process of creating complex, no so easy for understanding indexes. Observation of government (public administration) as a network of public bodies and institutions that works in direction of public service realization opens possibility of using new tools in the process of indicators creating. Social network analysis as a tool that examines the structure of relationships between social entities seems as appropriate for measuring e-government maturity by using public administration bodies (as a nodes) and information flow in the back office during the process of service delivery (as relation). Some of the network properties measured by social network analysis can be used as appropriate quantitative indicator for measuring e-government development. The findings of our research realized on real case studies in Republic of Macedonia confirms those assumptions and open opportunities for using social network analysis in the process of real-time evaluation and benchmarking of e-government maturity.

Keywords. e-government, e-government benchmarking, indicators, social network analysis.

1. Introduction

The ‘guiding star’ of governments in the process of introducing and implementation of e-government is more effective public administration tailored to the citizen needs. E-government as a concept can be defined as reform based on introducing information and communication technology (ICT) in public administration. ICT as a tool opens significant possibilities for improving quality of: a) external service (provision of public service delivery to citizens and business) well known as front office part of e-government; and 2) internal services (provision of service delivery between public administration institutions in the process of public services realization) well known as back office part of e-government.

The e-government implementation progress success is liable to measurement. Benchmarking as a method of measurement presents an ongoing process of continuous (mostly at yearly base) measurements and comparison of e-government development

process. Starting from 1998 to 2010 has been realized thirteen benchmarking studies [4]. All indicators and indexes used in the e-government benchmarking process have a focus on measuring front office: realizations of on-line services or creation and uses of web sites. Common for (almost) all indicators is that have been introduces with poorly design, by using simple statistical operations in the process of creating complex, and no so easy for understanding indexes. This is confirmed by the results of previous studies that point out: "no robust and well-defined measuring tool [are] available" [7]; Waksberg and Aibar [36] stress out that "missing network government indicators". The lack of well-defined network e-government indicators and indexes emphasize the need of establishing new indicators which background will not be only statistics methods.

The rest of this paper is organized as follows. Section 2 provides overview of e-government benchmarking indicators. Section 3 introduces social network analysis. The practical use of social network analysis is presented in Section 4. The results of using of social network analysis in the process of measuring back office integrity are presented in Section 5. Finally, Section 6 discusses and concludes the paper.

2. Situation with indicators in e-government benchmarking

The measuring on-line services and its sophistication; as well as creation and uses of web sites are dominate focus at the e-government benchmarking studies [1], [14], [35], [37]. In the process of e-government benchmarking are established a number of indicators by the researchers count more than 300 indicators [33].

The indicator is a pointer. It helps to understand complex realities by quantification of the phenomena. The indicator usually aggregates raw and/or processed data. The combination of few aggregated indicators is named indexes. Because of the longitudinal dimension of the benchmarking process, comparative of the indexes in the process of benchmarking are used to describe the performance of an institution, government, or whole country.

Table 1: Examples of e-government benchmarking index and indicators focused on front office measuring, and its estimation.

Author and year of publishing	**United Nations 2008**	**WASEDA 2011**	**Economist Intelligence Unit's 2010**
Number of indicators and indexes	3 indicators 12 sub-indicators 1 index	7 indicators 31 sub-indic 1 index	6 indicators 39 sub-indicators 1 index
Indicators	1. Web Presence Index **(WPI)** 2. Telecommunication Infrastructure Index **(TII)** 3. Human Capital Index **(HCI)**	1. Network prepare-ness - **NP** 2. Management optimization – **MO** 3. Required interface-functioning applications - **RIFA** 4. National portal- **NP** 5. CIO in government - **CIOG** 6. e-Government promotion – **eGP** 7. e-Participation - **eP**	1. Connectivity and technology infrastructure – **CTI** (20%) 2. Business environment - **BE** (15%) 3. Social and cultural environment – **SCE** (15%) 4. Legal environment – **LE** (10%) 5. Government policy and vision – **GPV** (15%) 6. Consumer and business adoption – **CBA** (25%)
	WPI (1. Official web	**NP -** Internet Users	**CTI** = (Broadband penetration (15%)+

Sub-indicators	site, 2. Type of available service delivery, 3. Presence of different type of services at different fields, 4. Use of portals, 5. Use e-government teams) **TII** (1. PSs/100, 2. Internet host/100, 3. % of population online, 4. Line/100, 5. Telecommunication) **HCI** (1. Human development index, 2. Info access index and urban % of total population)	+ Broadband Subscribers + Mobile Cellular Subscribers + PC Users **MO** - Optimization Awareness + Integrated Enterprise Architecture + Administrative and Budgetary Systems **RIFA** - Cyber Laws + e-Tender systems + e-Tax system + e-Payment system + e-Voting system + Social Security Service + Civil Registration + Consular Services + Labor Related Service + e-Health system **NP –** Navigation + Interactivity + Interface + Technical **CIOG** - GCIO Presence + GCIO Mandate + CIO Organizations + CIO Development Programs **eGP** - Legal Mechanism + Enabling Mechanism + Support Mechanism + Assessment Mechanism **eP** - e-Information and Mechanisms + Consultation + Decision-Making	broadband quality (10%) + broadband affordability (10%) + mobile-phone penetration (15%)+ mobile quality (10%) + Internet user penetration (15%) + international Internet bandwidth (10%) + Internet security (15%). **BE** = Overall political environment + macroeconomic environment + market opportunities+ policy towards private enterprise+ foreign investment policy + foreign trade and exchange regimes + tax regime+ financing + the labor market. (All nine criteria are weighted equally.) **SCE** = Educational level (measured by school life expectancy, gross enrolment in education and enrolment in tertiary education) + Internet literacy + degree of entrepreneurship+ technical skills of workforce + degree of innovation (measured by the generation of patents and trademarks, as well as R&D spending). (All five criteria are weighted equally.) **LE** = Effectiveness of traditional legal framework (30%) + laws covering the Internet (25%) + level of censorship (10%) + ease of registering a new business (25%) + electronic ID (10%). **GPV** = Government spend on ICT as a proportion of GDP (5%) + digital development strategy (25%) + e-government strategy (20%) + online procurement (5%) + availability of online public services for citizens (15%) + businesses (15%) + e-participation (15%, based on the UN e-participation index). **CBA** = Consumer spending on ICT per head (15%) + level of e-business development (10%) + use of Internet by consumers (25%, assessing both the range of Internet features used by individuals and their online purchasing activity) + use of online public services by citizens (25%) + businesses (25%).
Name of Overall index	**E-GOVERNMENT INDEX (EGI)**	**2011 WASEDA UNIVERSITY WORLD e-GOVERNMENT RANKING**	**DIGITAL ECONOMY RANKINGS 2010 (DER)**
Estimation of Overall index	**EGI=WPI+(TIIxHCI)**	**Not clear how is estimation of the final index**	**DER=CTI+BE+SCE+LE+GPV+CBA 100= 20%+15%+15%+10%+15%+25%**

The benchmarking process usually use overall index composed by several lower-level indicators and many sub-indicators that are estimated independently. Table 1 present three overall indexes of three benchmarking studies [14], [33], [35]. At the columns are presented the studies. At the rows are presented data for the studies as follows: at the first row are given references to the studies: the name of the study/author of the study and the period in which the indicators/indexes are used; the second row provides the number of indicators,

sub-indicators and indexes used in certain study as a base for benchmarking; the next two rows are related to the content of the indicators and sub-indicators, its names and the aggregation formula for its calculating (if is given by the authors). Finally, the last two rows are related to the final index of the study-base for benchmarking: first is given the full name of the index and than, the last row in the table gives formulas its estimation (if is given and easy recognized in the text of the study).

Each indicator has own methodology of creation. At the present [1], [14], [30], indicators used in the process of e-government benchmarking preferred methodological process which as backbone has web search and mail search - questionnaire (in an electronic form or mail), and sometimes face-to-face interviews. Document analysis or statistical data analysis as well as row data obtained from own research are the mostly used methods. Here all estimations are based on the data analysis tools: applied probability, sampling, estimation, hypothesis testing, linear regression, analysis of variance, categorical data analysis, and nonparametric statistics [19].

At the benchmarking studies, quantitative results usually are supported with comments by the authors of the study, what gives qualitative dimension. As the focus of benchmarking studies was front office of the e-government the results of this measurements "provides assets of useful information from a user-perspective, but does not provide any information how well the back office of e-government is organized" [17].

"The work on back-office measurement so far was limited" note UN [31]. There are few efforts to measure back office development [4] but all of them remain only at the level of evaluation. Some efforts for creating indicators which measured the G2G aspect in the process of e-government benchmarking were made by some authors [1], [30]. Concentrating on transforming government, back-office changing, connectivity and infrastructure, connectivity and network preparedness were some of the indicators. The hard and mostly expensive process of data collection as well as mostly qualitative description of the results, with out possibility of any comparative over the time, make this efforts unusable on long term.

3. Social Network Analysis (SNA)

Social network analysis (SNA) examines the structure of relationships between social entities. It is concerned with understanding the linkages among social entities and the implications of these linkages [34]. SNA analyze social relations in the network by using "a set of mathematical methods and distinctive methodology that encompassing specific techniques for collecting data, statistical analysis and visual representation" [23]. So, SNA has become a powerful methodological tool alongside statistics with main goal of discovery and elaborating patterns of social ties/relations among nodes/actors.

The key concept in SNA is the term "network" which present "a set of actors connected by a set of ties" [10] or speaking with SNA terminology it is a set of nodes connected with relation or set of relations. As nodes/actors in the social network can be single entities (people/persons, computers, concepts, URLs) or collective social units (group of people in a society, departments within a corporation, public service agencies, or nation/state). "Ties are connected pairs of actors" [8] representing relations established between nodes. Ties are seen as "pipes or roads along which traffic flows" [9]. The type and range of ties can be different: collaborations, friendships, web links, citations, trade ties, exchange of social support, resource flows, information flows, etc. [34].

The visualization of the connections between the actors presents graph. Mathematical theory of graphs allows formally defining and observing a range of network properties [26].

Network properties are generally classified in two groups: properties of the whole network and properties of the individual network nodes. Centrality refers to positions of individual nodes/actors in the networks, whereas the centralization characterizes the entire network. In this sense, a central node in the network is the one which is connected through fewer links to every other node in the same network. On the other hand, when studying the centralization of a network - a network with one central node is more centralized as compared to a network with various equalized centrality measures [23].

4. Practical use of SNA

Many studies which focus are inter-organizational networks employ social network analysis as a methodological tool [25]. The studies have different focus: coordination of soft-target organizations [29]; analysis of partnership networks for implementing an e-government project [11], [12], implementing local e-government policy [20]; TV production projects [28], and video-games sector [32]; film industry [18]; inter-organizational networks in the domains of wood-products manufacturing industry [16]; development of trauma [3]; analysis of health-care networks induced from the path that the patient makes through different hospital departments [2]; as well as networks related to the delivery of health services related to chronic mental diseases [15], [21], [22], [24].

E-government and public administration is not often in the focus of the studies that use SNA [25], with exceptions of few studies [11], [12], [21], [29]. Morrissey [21] use approach which combines perspectives from inter-organizational theory and the techniques of social network analysis. Its work mainly has been focused on how procedures can be used to assess coordination polices in the process of service delivery protocol creation.

Starting of the idea presented by Alter and Hage [2] on analysis of health-care networks, become idea for use SNA in process of back office measurement [5]. In fact, similarly as the path that the patient makes through different hospital departments, which compares the properties of these networks to the efficiency thereof, the idea is to follow the path of citizen or business do during the process of service delivery while the citizen or business provide necessary documentation for that public service realization. The idea is instead analysis of health-care networks, to analyze service delivery protocols; and instead measuring path of the patient to analyze document-flow between institutions in the process of service delivery.

This approach is base to start to think about applying social network analysis methods to inter-organizational networks in public administration where the actors are public administration bodies and institutions while the ties present information flow during the process of service delivery.

5. The results of using of social network analysis in the process of measuring back office integrity

In the process of search indicator(s) that can be used in the process of e-government benchmarking and which focus will be back office, Bogdanoska Jovanovska [4] starts with analyzing the inter-organizational networks in public administration established through delivery of public services. Those network they named as “document-flow networks” based on the tie that connect actors in the network. Those networks include actors of three different types: citizen, public administration institutions and other institutions (businesses and others) that the citizen need to obtain all the documents related to the service of interest.

Because of the longitudinal dimension of benchmarking process as well as to be in position to compare the results, it was necessary of few separate consecutive measurements. For that reason it was necessary to have different networks for the same service delivery. The first network was created base on actual service delivery protocol for separate public service in each public administration area. The second (for comparative purpose created) network for each under-researched service was created: based on simple simulation of back office integration (2010) and law reforms in service delivery procedures (2012). All together five public administration areas in Republic of Macedonia were topic of the research: urbanism (2010), taxation, state and social pensions, civil registry and business registration (2012) with analyze of more than 40 complex public services that belongs in those public administration areas.

The expectations are: the integration of the public administration institutions back office will cause changes in network properties in direction of increasing of centrality of entire network as well as increasing of centrality of public administration institutions as a node, while the centrality of the citizen as a node will decrease. In situation to receive those results, we can conclude that SNA can be useful tool for measuring e-government maturity.

6. Conclusion

The starting point for the idea of using social network analysis as a method for research e-government development by back office perspective first of all was awareness that states institutions and bodies are part of public administration as a network and second, knowledge that the difference between "functional-oriented" government and 'citizen-oriented" government lie in back office integration.

Traditional way of public administration is organized following the function-oriented 'silo effect' scheme where the citizen has to go from one organization to another in order to establish a proper document-flow between public administration institutions that lead to service realization. This point out at low level of (e)government development i.e. low level of back office integration what is related to frequent activities of the citizen in the process of realization of the complex service. In contrast, mature public administration and e-government services, supported with appropriate integrated back-office systems, have focus on public service delivery from the perspective of the citizen, so called citizen-centric e-government, where citizen usually gets all the needed services in a 'one-stop-shop' front-office. Thus, the PA institutions take active role in the document-flow network and ease the burden of citizen in the processes of service delivery and realization [4].

The results in both attempts [4] confirm the expectations:

a) Centrality measure at level of individual node:

- Centrality of the "citizen" node decreased (as was expected). This is due to the fact that in mature service-deliver environments, the citizen has a peripheral role in the process of service delivery and most of the document flow is independent from the citizen intervention. In contrary, underdeveloped service delivery environment is characterized with high centrality of the citizen node in the document-flow network (2010).
- Centrality for the nodes corresponding to PA institutions shows that back-office development move PA institutions from the periphery of the document-flow network towards the center. The possible explanation is that in mature service-delivery environment with strong back-office support many PA institutions play a

role of a proxy or intermediate that delivers a small part of the service, or only establish a connection between other institutions.

b) Centralization at the level of whole network:

- Centralization of the whole network decreases. This decreasing is mostly due to the reduced centrality of the citizen node, so there is more than one central node.

This network properties measured by social network analysis can be used as appropriate quantitative indicator for measuring e-government development.

References

1. Accenture, Government Executive Series, Leadership in Customer Service: Delivering on the Promise, High Performance, Delivered, Accenture, 2007.
2. Alter C., J. Hage, *Organizations Working Together***,** Sage publications, International Educational and Professional Publisher, Newbury Park, London, New Delhi, 1993.
3. Bazzoli J.G., R. Harmata, C. Chan, *Community-based trauma systems in the United States: an examination of structural development*, Soc. Sci. Med., Vol. 46, No. 9, pp. 1137-1149, 1998.
4. Mimoza Bogdanoska Jovanovska, *Evaluating e-Government through Analysis of Inter-Organizational Networks in Public Administration*, doctoral dissertation, Faculty of Administration, University of Ljubljana, 2012.
5. Bogdanoska Jovanovska M., Lj. Todorovski, *A Comparative Review of Three Back Office Benchmarking Studies*, The 19th NISPAcee Annual Conference 2011, Varna, Bulgaria, 2012.
6. Bogdanoska Jovanovska M., N. Erman, Lj. Todorovski, *Indicators of intensity and development of e-government back-office. Indicators models, metrics and approaches for monitoring and evaluation of the situation in the public sector*, monograph, Faculty of Administration, University of Ljubljana, 2012.
7. Bongers F., C. Holland, R. Vandeberg, *Measuring E-Government*, IADIS International Conference e-Society, pp. 195-203, 2003.
8. Borgatti P.S., C.P. Foster, *The Network Paradigm in Organizational Research: A Review and Typology*, Journal of Management, Vol. 29 (6), pp. 991–1013, 2003.
9. Borgatti P.S., H. Li, *On Social Network Analysis in a Supply Chain Context*, Journal of Supply Chain Management, Vol. 45(2), pp: 5-22, 2009.
10. Borgatti P.S., S.D. Halgin, *On Network Theory*, Organization Science, Vol. 22(5), pp. 1168-1181, 2011.
11. Cotterill S., S. King, *The Role of Social Networks in the Development of English Local E-Government*, Political Studies Association Conference, 2006.
12. Cotterill S., S. King, *Public Sector Partnerships to Deliver Local E-Government: A Social Network Study*, In: Wimmer, M.A., Scholl, H.J., Grönlund, A. (eds.) EGOV 2007. LNCS, Vol. 4656, pp. 240-251, Springer, Heidelberg, 2007.
13. Dutta S., I. Mia, *The Global Information Technology Report 2010–2011: Transformations 2.0*. World Economic Forum and INSED, 2011.
14. EIU - Economist Intelligence Unit. Digital economy rankings 2010, Beyond e-readiness. A report from the Economist Intelligence Unit, written in co-operation with The IBM Institute for Business Value. Economist Intelligence Unit Limited 2010, 2010.
15. Fried B., M. Johnsen, B. Starrett, M. Calloway, J. Morrissey, *An empirical assessment of rural community support networks for individuals with severe mental disorders***,** Community Mental Health Journal**,** New York, Vol. 34, Iss. 1, pg. 39, 1998.
16. Human E.S., G.K. Provan, *Legitimacy Building in the Evolution of Small-Firm Multilateral Networks: A Comparative study of Success and Demise*, Administrative Science Quarterly, Vol. 45 pp. 327-365, Johansen et al 1996, 2000.
17. Janssen, M. (2010). Measuring and Benchmarking the Back-end of E-Government: A Participative Self-assessment Approach, M.A. Wimmer et al. (Eds.): EGOV 2010, LNCS 6228, pp. 156–167, IFIP International Federation for Information Processing 2010
18. Krätke S., Network analysis of production Clusters: the Potsdam/Babelsberg Film ndustry as an

example, *European Planning Studies*, Vol. 10, No. 1, 2002.
19. Manevska V., *Voved vo teorijata na verojatnost i matematička statistika*, Tehnički Fakultet Bitola, MK. ISBN 998978633X, 9789989786334, 2003.
20. Medaglia R., *Local Networking for e-Services: A UK Case Study*, In: Wimmer, M.A. et al. (eds.) EGOV 2006. LNCS, vol. 4084, pp. 256-268, Springer, Heidelberg, 2006.
21. Morrissey P.J., M. Tausig, L.M. Linsey, *Community Mental Health Delivery Systems. A network perspective*, The American Behavioral Scientist (pre-1986), 1985, 28,5ABI/INFORM Global pg. 704, 1986.
22. Morrissey P.J., M. Calloway, W.T. Bartko, M.S. Ridgely, H.H. Goldman, I.R. Paulson, *Local Mental Health Authorities and Service System Change: Evidence from the Robert Wood Johnson Foundation Program on Chronic Mental Illness*, The Milbank Quarterly, Vol. 72, No. 1, pp. 49-80, Blackwell Publishing on behalf of Milbank Memorial Fund, 1994.
23. Nooy D.W., A. Mrvar, V. Batagelj, *Exploratory social network analysis with Pajek*, Cambridge University Press, New York Provan 2007, 2005.
24. Provan K., H.B. Milward, *A Preliminary Theory of Interorganizational Network Effectiveness: A Comparative Study of Four Community Mental Health Systems*, Administrative Science Quarterly, Vol. 40, No 1, pp 1-33, 1995.
25. Provan, G. Keith, Fish, Amy and Sydow, Joerg (2007). *Interorganizational networks at the Network level: A review of the Empirical literature on Whole Networks*, Journal of Management, Vol. 33, pp. 479.
26. Scott J., *Social Network Analysis*, Handbook, SAGE Publication, 2000.
27. Snellen I., *Matching ICT Networks and PA Networks: Lessons to be learned*, IIAS Governing Networks, EGPA Yearbook, IOS Press, Amsterdam, pp. 129 – 142, 2003.
28. Soda G., A. Usai, A. Zaheer, *Network Memory: The Influence of Past and Current Networks on Performance*, Academy of Management Journal 2004, Vol. 47, No. 6, pp: 893–906, 2004.
29. Uddin M.S., L. Hossain, *Towards coordination preparedness of soft-target organisation*, In: Wimmer, M.A., Scholl, H.J., Janssen, M., Traunmüller, R. (eds.) EGOV 2009. LNCS, Vol. 5693, pp. 54-64, Springer, Heidelberg, 2009.
30. UN - United Nations, *UN Global E-government Survey 2008: From e-government to Connected Governance*, Department of Economic and Social Affairs Division for Public Administration and Development Management, New York, 2008, (http://unpan1.un.org/intradoc/groups/public/ documents /un/unpan028607.pdf)
31. UN DESA, Expert Group Meeting E-Government Survey: Getting to the Next Level, United Nations, New York, 2009.
32. Venkatraman N., C. Lee, *Preferential linkage and network evolution: a conceptual model and empirical test in the U.S. video game sector*, Academy of Management Journal, Vol. 47, No. 6, pp. 876-892, 2004.
33. Vintar M., M. Kunstelj, A. Leben, *Benchmarking the Quality of Slovenian Life-Event Portals*, Improving the Quality of East and West European Public Services, E. Loffler and M.Vintar (eds) Ashgate, Hampshire, pp. 208–221, 2004.
34. Wasserman S., K. Faust, *Social Network Analysis: Methods and Applications*, University Press, Cambridge, UK, 1994.
35. Waseda, The 2011 Waseda University World e-Government Ranking, WASEDA, 2011.
36. Waksberg A., E. Aibar, *Towards a Network Government? A Critical Analysis of Current Assessment Methods for E-Government*, Lecture Notes in Computer Science, Vol. 4656, pp. 330-341, 2007.
37. World Economic Forum and others (INSEAD, infoDev). *The Global Information Technology Report 2010–2011, Transformations 2.0*, 10th Anniversary Edition, 2011.

Identification of Clusters from Data Sets

Vesna Mufa[1], Violeta Manevska[2]

[1]St. Kliment Ohridski University, Faculty of Administration
and Information Systems Management,
Partizanska bb, MK-7000 Bitola, Macedonia
vesna_mufa@hotmail.com,

[2]St. Kliment Ohridski University, Faculty of Administration
and Information Systems Management,
Partizanska bb, MK-7000 Bitola, Macedonia
violeta.manevska@uklo.edu.mk

Abstract. The era in which we live is called information age, i.e. the era of data. The evolution of data mining began with the storage of data in a computer. Data mining is analyzing of large data sets in order to find unexpected relationships and summaries of data on new, previously unknown, comprehensible and useful ways. One of the most common tasks of data mining is a descriptive modeling, which describes data and processes that they generate. One of the most used descriptive methods is a clustering. The clustering aims to detect natural groups (clusters) into the data. The purpose of this paper is the identification of clusters, their presentation and visualization. The clusters are identified through partitioning, as the basic version of clustering, which divides the instances of several mutually exclusive clusters, with using of the Euclidian distance. Determination of the exact number of clusters is important because the correct number of clusters is not only an input parameter, but it controls the granularity in cluster analysis too.

Key words: Data Mining, Descriptive Modeling, Clustering, Clusters

1. Introduction

Living in the information age, we come across with data that are stored on a computer, with a belief that information leads to success. In this way, the amount of data increases every year more and more, but there is less and less useful information. Manual analysis of data in order to extract useful information, is an impossible mission. An option is the use of computer programs that enable the analysis process.

Data mining is analyzing of large sets of data in order finding unexpected relationships and summaries of data on new, previously unknown, comprehensible and useful ways [1].

The evolution of data mining began with storing data in a computer. The evolutionary process of data mining ranges from retrospective access to data to prospective and proactive delivery. Data sets reach sizes of gigabytes and terabytes that result in obtaining interesting, understandable and reliable patterns and summaries.

One of the most common tasks of data mining is a descriptive modeling, which describes data and processes that they generate [2]. The descriptive models discover patterns or trends in data, which should be as transparent as possible, which implies that the results

from the models should be appropriate to the intuitive explanation and interpretation. The best known descriptive methods are: clustering and association rules.

2. Clustering

Clustering or cluster analysis is an automatic classification of the instances into groups, so that the instances within a group are as similar as possible, while the instances from different groups are as different as possible [3]. The formed groups are called clusters. The cluster analysis can be specified as an optimization problem, where the inter-cluster similarity and intra-cluster diversity should be optimized. Due to the difficulty of its realization, the clustering methods try to find an optimal solution [4].

In the process of clustering the input is a data set and a measure of similarity between two instances, and the outputs are the clusters. The clusters can be described by using predictive models such as the rules and trees, which is essential for a deeper analysis of the characteristics of the data [5]. Clustering serves as a preparatory step for the other methods such as the characterization and classification that are performed on the detected clusters.

The requirements for clustering are:

- **Scalability:** The algorithms should be able to form clusters of small and large data sets.
- **Ability to deal with different types of data:** Although the most appropriate type is continuous data, the algorithms need to deal with the other types of data, such as a binary, nominal and ordinal data, and their combinations. Also, shouldn't be left out the complex data types, such as the graphs, sequences and images.
- **Discovering clusters with an arbitrary shape:** Many algorithms use the Euclidean or Manhattan distance for determination of clusters with a spherical shape, which have similar size and density. Because the cluster should have an arbitrary form, this imposes the use of algorithms that form clusters with arbitrary shape.
- **Domain knowledge for determining the input parameters:** The input parameters are difficult to determine when the data set consists of many attributes and when users don't understand the data. Therefore, domain knowledge is necessary for setting the input parameters, such as the desired number of clusters and the similarity measure.
- **Ability to deal with incorrect data:** This situation occurs if the cleaning of data is omitted and results in generated clusters with undesirable quality. This situation is overcome by the robustness of the algorithms.
- **Incremental clustering:** It allows incorporation of incremental updates (presentation of new data) into existing clusters at any time.
- **Capability for clustering data with many attributes (dimensions):** Finding clusters from data with many dimensions is a challenging task.
- **Interpretability:** Clustering should produce understandable and useful results. It is important to assess how the objectives, which should be achieved, influence on selection of the input parameters and methods of clustering [6].

The different types of clustering differ each other in terms of the level of partitioning, whether the clusters are mutually exclusive and in terms of the measures of similarity that they use. According to the level of partitioning, all obtained clusters can be on the same conceptual level, without any hierarchy between them or can be hierarchically subordinate. In some situations, the instances belong to only one cluster, which means that the clusters are mutually exclusive, but in some situations, the instances belong to more clusters, which leads to the destruction of the rule of mutual exclusion. The similarity measures are different, so some of the algorithms use distance as a measure of similarity, while some use the density.

3. Application of the Clustering on Real Data

The data set, which is a subject to clusters identification, is composed of components of a hard cheese produced in a traditional manner by two manufacturers (M1 and M2), components of a hard cheese produced on an industrial way by a manufacturer (IM), and a phase in which the hard cheese is, which have a role of attributes. The values of the components are obtained by laboratory tests. From each manufacturer are taken 10 samples that are examined after completion of the various phases (coagulum, salting and ripening), and after 20 days and 45 days after hard cheese is placed in brine. In Table 1 are given the names of the attributes and their type [7].

Table 1. Names and types of attributes

Names of attribute	Types of attribute
ash	continuous
fats	continuous
salt	continuous
moisture	continuous
dry_substance	continuous
total_nitrogen	continuous
total_proteins	continuous
coefficient_of_maturity	continuous
pH	continuous
SH	continuous
phase	discrete {coagulum, after_ripening, after_salting, 20_day, 45_day}
manufacturer	discrete {M1, M2, IM}

The clustering is performed by the implementation of partitioning, i.e. by using the **SimpleKMean** algorithm and the Euclidean distance.

Partitioning is the simplest, basic version of clustering that shares the instances of mutually exclusive clusters. The number of clusters that the algorithm will generate is specified by the user and is a starting point. Execution of the algorithm multiple times results in selection of number of clusters as the most appropriate solution. Besides the number of clusters, should be stated and the initial states of the clusters.

The goal of this type of clustering is minimizing variability within the clusters and maximizing variability between the clusters. If data are presented in Euclid space, the members of each cluster belong, or are instances of the entire set. The intersection of the clusters is an empty set.

The best-known partitioning algorithm is the k-means algorithm, which is a centroid based. The centroid is defined as a mean value of the points that form the cluster. Initially, randomly are select k instances of data set, which represents the centers of the clusters. The other instances are accompanied to the most similar clusters, based on the smallest Euclidian distance between each instance and the center of the cluster. The algorithm improves the variation within the cluster by calculating the new center for each cluster based on the

assigned instances of the previous iteration. Iterations end when the clusters formed in the last iteration is equivalent to the clusters formed in the previous iteration.

Due to the use of the mean value of the clusters, the algorithm restricts the implementation of discrete attributes. The modifications aim to form clusters of discrete data. The original and modified version differs in terms of the similarity measure and from the method of calculating the center of the cluster. One such modified version is k-modes, where the means are replaced with the modes. If a data set consists of continuous and discrete attributes, the clusters are formed with an integrated use of the k-means and k-modes algorithms.

In the most clustering algorithms, the similarity between instances is represented by calculating the distance between each pair of instances. There are a number of methods for calculating the distance, but none of them is the best for general use. Some of them calculate the distance between instances whose attributes have equal importance (Manhattan distance and Chebyshev distance), while some take into account the importance of attributes, represented by their weights (Minkovski distance and measure for distance with mixing the weights). The equation for calculating the Euclidian distance adjusts for both cases.

The distance between the two instances is marked with $dist\ (A, B)$, and the instances are vectors $A = (a_1, a_2, \dots, a_n)$ and $B = (b_1, b_2, \dots, b_n)$.

The Euclidian distance uses the Pythagorean Theorem. The length of the segment that connects the point $A(a_1, a_2)$ that is an instance of the training set, with the point $B(b_1, b_2)$ that is an instance of classification is calculated by the equation:

$$dist(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}. \quad (1)$$

The Euclidian distance between points $A(a_1, a_2, \dots, a_n)$ and $B(b_1, b_2, \dots, b_n)$ in n-dimensional space is:

$$dist(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}. \quad (2)$$

Due to the measurement of various attributes in terms of different scales, the direct implementation of the Euclidian distance introduces the effect of the drastic difference between the attributes [8]. Also, if there are big attribute values, the distance calculation is more complicated in contrast from attribute with smaller values. For these reasons, is applied normalization of the values of the attributes in order to belong to the interval $[0,1]$. The normalized value is calculated according to this equation:

$$a_i = \frac{v_i - min_{v_i}}{max_{v_i} - min_{v_i}}, \quad (3)$$

where v_i is the value of the attribute i, max_{v_i} and min_{v_i} are the maximum and minimum value of the attribute i. The weights of the attributes are introduced for highlighting the priority of certain attributes over others. Using of weights when measuring the distance is achieved through the transformation of the Euclidian distance equation, which is:

$$dist(A, B) = \sqrt{w_1(a_1 - b_1)^2 + w_2(a_2 - b_2)^2 + \cdots + w_n(a_n - b_n)^2}, \quad (4)$$

where $w_1, w_2 \dots, w_n$ are appropriate weight. The values of the weights are scaled for their sum to be equal to one [9].

The nature of the data set requires to be used a combination of k-means and k-modes. The algorithm is adjusted for getting five clusters that are shown in Figure 1. The descriptive model displays the centroid of all data and the centroids of formed clusters. The model

shows the number of instances that belong to each cluster. To the first cluster belongs 19% of the data, 20% in the second and the third, 15% in the fourth and 27% to the fifth cluster. From received clusters can be observed that different values of the attribute *phase* are contained in the centroids of five clusters. The same refers to the attribute *manufacturer*, with a difference that there is redundancy due to the existence of three different values of the attribute and five clusters.

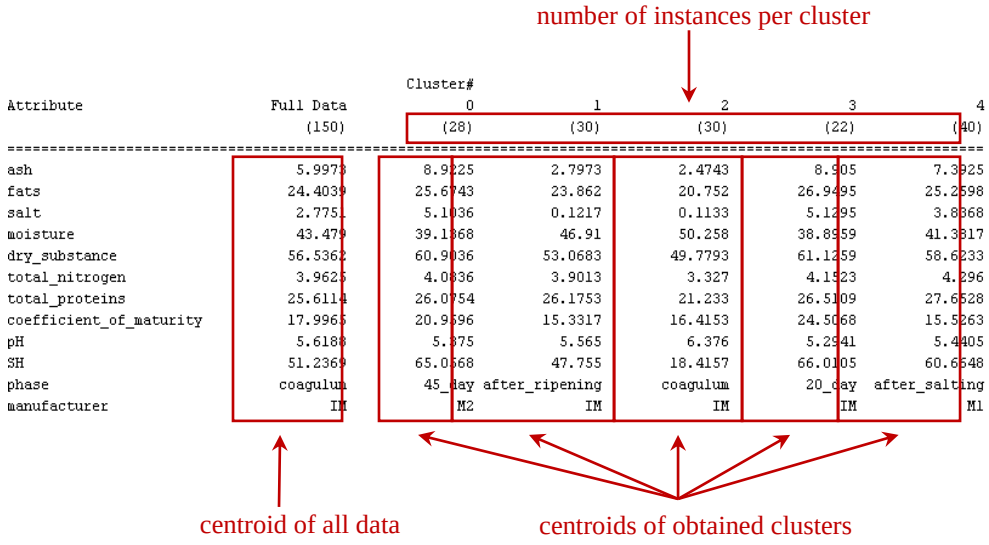


Figure 1. Descriptive model

Figure 2 shows a visual representation of the obtained clusters. The x-axis present the different phases, while to the y-axis is assigned the coefficient of maturity. Attributes in terms of the axes can be altered, so the obtained different clusters depend from selected attributes.

Figure 2. Visual display of the obtained clusters

Determination of which cluster will belong the new instance is done by calculating the Euclidian distance from the new instance to the centroids of clusters. The instance is assigned to the cluster whose distance between the instance and the cluster is minimal.

4. Determination of Appropriate Number of Clusters

Determining the exact number of clusters is important because the exact number of clusters is not only an input parameter, but it controls the granularity of cluster analysis. It means finding a balance between accuracy and compressibility. If the entire data set represents a cluster, then the compression of data is maximized, but the clustering has no value. On the other hand, if each instance represents a cluster, then the accuracy increases, but prevents the summarization of data.

Determination of the number of clusters is not an easy task because the appropriate number of clusters is mostly random. It usually depends upon the distribution of data and their form. There are several ways to calculate the appropriate number of clusters.

The first way, which is the simplest, is the number of clusters to be equal to $\sqrt{(n/2)}$, where n is a number of instances. The expected number of instances for each cluster is approximately $\sqrt{2n}$.

The second way means that with increasing the number of clusters is reduced the sum of intra-cluster variation. The existence of multiple clusters contributes to the discovery of groups whose instances are more similar to each other. However, the reduction of the sum of intra-cluster variation can lead to a large number of formed clusters. So as heuristics, the sum of intra-cluster variation and the number of clusters are represented in two-dimensional space in the form of a curve, where the number of clusters of transition point is chosen as the appropriate number of clusters.

Performing the same approach for clustering multiple times for different values of k (number of clusters) is the latest way for determining of the appropriate number of clusters. The obtained clusters are evaluated by the user who has a domain knowledge or has information for an ideal clustering. From all performances, the number of clusters that best divides data is selected as an appropriate number of clusters.

5. Conclusion

The data set for hard cheese is a suitable example of a practical illustration of clustering as a method of descriptive modeling. The clustering is performed with implementation of partitioning, i.e. by using the SimpleKMean algorithm and the Euclidian distance.

According to the level of partitioning, all obtained clusters are on the same conceptual level. Determination of which cluster will belong the new instance is done by calculating the Euclidian distance from the new instance to the centroids of the clusters. The instance is assigned to the cluster whose distance between the instance and the cluster is minimal. Determination of the appropriate number of clusters is a balance between the compressibility and the accuracy of the resulting model.

References

1. Hand D., H. Mannila, P. Smyth, *Principles of Data Mining*, Massachusetts Institute of Technology, London, UK., 2001.
2. Mufa V., V. Manevska, B. Nestoroska, *The Data Mining Process*, 7th International Conference for Information Systems & Grid Technologies, Sofia, Bulgaria, 2013.
3. Nestoroska B., V. Manevska, V. Mufa, *Marketing Research by Applying the Data Mining Tools*, 7th International Conference for Information Systems & Grid Technologies, Sofia, Bulgaria, 2013.
4. Gupta G. K., *An Introduction to Data Mining with Case Studie*s, Monash University, Clayton, Australia, 2006.
5. Kantardzic M., *Data Mining: Concepts, Models, Methods and Algorithms*, A John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2011.
6. Han J., M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, Elsevier Inc., San Francisco, USA, 2012.
7. Mufa V., *Information Extraction with Implementation of Data Mining Methods*, master thesis, FAMIS – Bitola, Macedonia, 2013.
8. Bramer M., *Principles of Data Mining*, Springer, London, UK, 2007.
9. Witten I. H., E. Frank, M. A. Hall, *Data Mining: Practical Machine Learning, Tools and Technique*, Morgan Kaufmann Publishers, San Francisco, USA, 2011.

Creating Annotations and Indexes in System for Intelligent Tagging and Search in Unstructured Data

Maria Marzovanova[1]

[1] UNWE, 1700 Sofia, Bulgaria
mmarzovanova@unwe.bg

Absract. The system for intelligent tagging and search uses an extended taxonomic structure as a basis for all the processes concerning tagging and search. The developed system has two very specific features – automated tagging of the unstructured data and intelligent tagged search. The focus of the current report is on the logic and the actual implementation behind these main processes. Each step important for the proper working of the system is illustrated in the terms of a small example which includes a part of a taxonomy and some text files with appropriate for the sample taxomy's content.

Keywords. Tagging, search, GATE, unstructured data, annotations

1. Introduction

One of the popular and well-known approaches for working with unstructured data is categorization and the developed solutions concerning this problem tend to match the available structured data with the categories created for the unstructured and in this way creating some order out of the mess. It's a process of identifying named and recognizable elements, such as names, places, etc.

The methodology implemented in the presented system for intelligent tagging and search is based on five methods for intelligent tagging and search using taxonomies and folksonomies. It can be classified as one of the approaches mentioned above but it should be mentioned that the unstructured data is categorized automatically based on a list of terms taken out of a taxonomy and the matching afterwards is done according to rules taken out of the taxonomy structure, the relations defined in it.

This report presents one implementation of the methodology with illustrations of the results after each step using a sample taxonomy and text files that need to be searched.

2. Decisions for tagging and indexing

Methods for tagging and indexing used in this system's have been presented in previous reports [1], [2] . So here I will include only a reminder of the whole picture in order to illustrate part of it with a working example, explain the mechanism behind the system and draw some ideas for further development.

One of the main problems that the system's development confronted was the wide variety of file formats in which unstructured data exists, each having its own structure and representation specifics. The differences between the file formats influence the way systems access and process their content. As the system for intelligent tagging and search needs to

treat the different unstructured-data sources equally, the solution is to develop mechanism for format recognition and separate interfaces for each of these formats. Text engineering is the issues's study area, and a number of software tools and development environments are available to help. Among the popular software dealing with the problem GATE is known to be the most appropriate for the needs of the system [8] [9].

GATE has been designed and developed for annotating documents, and all the functionality it offers is related to that. It works with language and processing resources. The presented system uses a corpus of documents, where all files with unstructured data are added, and an annotating component called ANNIE Gazetteer. ANNIE GAZETTEER is responsible for getting a list of terms, extracted in advance from a certain taxonomy/folksonomy, and annotating every occurrence of them in the content of the corpora documents. As a result, the process creates an XML file for each document that has all the content of the original source extended with tags placed around each occurrence of each term. The process returns also a list of all added annotations which is daved in the system's database and is used for creating indexes and pointing which term in which file can be found.

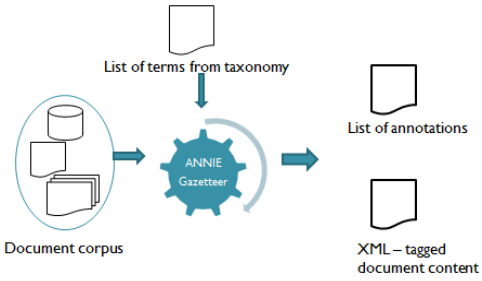


Figure 17

The idexes picked out of the annotations have a structure holding data for a connection between a term from a taxonomy/folksonomy and its occurrences in the content being searched. To identify the term with its specific use in a certain taxonomy it has the term's id and the taxonomy/folksonomy id. The document id and the annotation position constitute the second group of fields. The indexing as a process is entirely automated and starts right after the end of tagging.

The system is designed to provide a wide range of search options. Starting from the mechanisms closest to the database of terms, the first process is analyzing the user's query. The main purpose at this point is to recognize the terms saved in the system in the entered query. As a result all indexes containing the entered word/s or similar to it are visualized, so that the user can choose which one to proceed with. From the system's point of view this gives information about the taxonomy and the exact term use.

3. Taxonomy and text documents

The taxonomy that will be used is a small structure containing six terms, one on the top level and five equally related to it on the second level of the hierarchy. The parent term is stock and its child terms are markets, risk and capital, insurance and shares (Figure 18). With a taxonomy like this we can tag instructured data containing information in the finance area. This means that if we find in a text document anything related to three of these five terms the text is about „stock" in the aspect of finance.

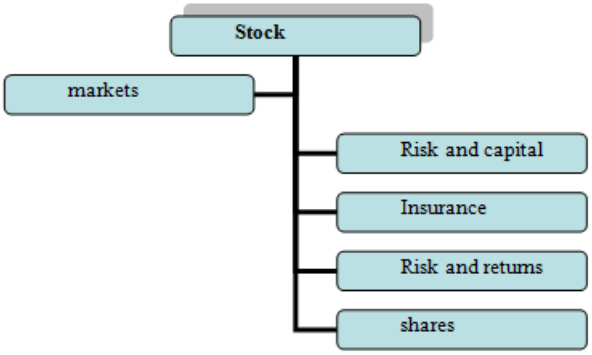


Figure 18

For testing purposes there are two files added to the system's repository. In order to demonstrate the system's ability to work with different file formats, they are chosen to be txt, and pdf. As the parent term in the taxonomy is "stock", this will be the term, for which one the files will be delivered as a result of eventual search using the system. Therefore, in terms of contents there is one file related to finances and one file containing the word stock with different meaning.

The first document is titled Stock market and is in txt format. It contains four of the terms used in the taxonomy branch of the term „stock" (Figure 19).

Stock market

A **stock market** or **equity market** is the aggregation of buyers and sellers (a loose network of economic transactions, not a physical facility or discrete entity) of stocks (shares); these are securities listed on a stock exchange as well as those only traded privately.....

Market participants include individual retail investors, institutional investors such as mutual funds, banks, insurance companies and hedge funds, and also publicly traded corporations trading in their own shares....

Figure 19

And the second document which is in pdf format is titled Product settings. One usages of the only term among the mentioned in the taxonomies, stock, is highlighted in **Figure 20**

Stock

Stock Status	Manually increment (by entering a positive quantity) or decrement (by entering a negative quantity) the stock quantity of a product.The quantity you enter (which can also be zero), will be automatically updated as this product is shipped or returned.....
Do Not Allow Back Orders	For products that can be back-ordered, stock status can reach negative values if the product is ordered after its stock value has reached zero...
Link Stock Status with ProductCode	Enter another product's code to have this product's stock status shared with that product to link product stock from one product to another. This is often used to offer a free accessory with purchase of one product so that the stock of one product never exceeds the stock of the free accessory.

Figure 20

There are two groups of actions according to the implemented in the system methodology that have to be performed before an actual search [4]:

- tagging the sources;
- indexing the tags.

These two methods are fully automated and require no user interaction. The automated execution is based on the logic lying in the taxonomy.

4. Tagging and annotations

The tagging process includes as a first step preparing a list of all the terms participating in the taxonomy. In this sample the system will create the following list:

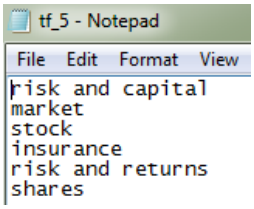


Figure 21

Once the list is full, the system process using GATE Embedded based application accesses the files added in the repository and starts the tagging process. Tagging as it is in the system provides two groups of results – new xml files with tagged sources' content and data for the added annotations for the database.

There is a small java application developed for this functionality. It uses GATE Embedded to create a corpus of Documents and a Gazetteer resource which is using the lists of terms as an Annotation set. During the execution of the application, the system creates xml documents for each of the repository files, containing the document's text and annotations corresponding to the terms from the lists.

The files created after performing tagging with the sample taxonomy over the two files we have added are shown in **Figure 22** and **Figure 23**. Each appearance of any item from the list is surrounded with a Lookup tag which indicates that the element is part of the terms list and a number of the annotation.

Stock market.txt.xml*

```
<paragraph xmlns:gate="http://www.gate.ac.uk" gate:gateId="0" gate:annotMaxId="9">Stock
<Lookup gate:gateId="1" majorType="terms">market</Lookup>
A <Lookup gate:gateId="2" majorType="terms">stock</Lookup>
  <Lookup gate:gateId="3" majorType="terms">market</Lookup> or equity
  <Lookup gate:gateId="4" majorType="terms">market</Lookup>
  is the aggregation of buyers and sellers (a loose network of economic
  transactions, not a physical facility or discrete entity) of stocks (<Lookup gate:gateId="5"
majorType="terms">shares</Lookup>); these are securities listed on a
<Lookup gate:gateId="6" majorType="terms">stock</Lookup> exchange as well as those
only traded privately.…
Market participants include individual retail investors, institutional
investors such as mutual funds, banks, <Lookup gate:gateId="7" majorType="terms">insurance</Lookup>
  companies and hedge funds, and also publicly traded corporations trading in their own
  <Lookup gate:gateId="8" majorType="terms">shares</Lookup>….
</paragraph>
```

Figure 22

Product settings.pdf.xml*

```
<html xmlns:gate="http://www.gate.ac.uk" gate:gateId="0" gate:annotMaxId="30"><head gate:gateId="1">
  <meta gate:gateId="2" content="1" name="xmpTPg:NPages"><meta gate:gateId="3" content="2014-01-19T16
<p gate:gateId="15">Stock
</p><p gate:gateId="16">Stock Status Manually increment (by entering a positive quantity) or decremen
entering a negative quantity) the <Lookup gate:gateId="23" majorType="terms">stock</Lookup> quantity
you enter (which can also be zero), will be automatically updated as this
product is shipped or returned.…
</p><p gate:gateId="17">Do Not Allow Back
Orders
</p><p gate:gateId="18">For products that can be back-ordered,
<Lookup gate:gateId="24" majorType="terms">stock</Lookup> status can reach negative
values if the product is ordered after its <Lookup gate:gateId="25" majorType="terms">stock</Lookup>
  value has reached zero...
</p><p gate:gateId="19">Link Stock Status
with ProductCode
</p><p gate:gateId="20">Enter another product&amp;apos;s code to have this product&amp;apos;s
<Lookup gate:gateId="26" majorType="terms">stock</Lookup> status shared with
that product to link product <Lookup gate:gateId="27" majorType="terms">stock</Lookup>
  from one product to another. This is often
used to offer a free accessory with purchase of one product so that the
<Lookup gate:gateId="28" majorType="terms">stock</Lookup> of one product never exceeds
the <Lookup gate:gateId="29" majorType="terms">stock</Lookup> of the free accessory.
</p><p gate:gateId="21">
</p><p gate:gateId="22"> </p></p></div></body>
</html>
```

Figure 23

Important data for the annotations added to the files is saved in the database. The database table holding that information has three meaningful columns – Annotation, DocumentUrl and StartNode. They contain all necessary marks to recognise each usage of a term in the unstructured data. The annotations added after tagging the sample documents are shown in Figure 24. What is important to notice there is that the system has found four different terms in Stock market.txt and only one in Product settings.pdf.

IDA...	Annotati...	DocumentUrl	StartNode
2001	market	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	1
2002	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	2
2003	market	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	3
2004	market	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	4
2005	shares	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	5
2006	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	6
2007	insurance	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	7
2008	shares	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Stock market.txt	8
2009	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	23
2010	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	25
2011	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	24
2012	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	27
2013	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	26
2014	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	29
2015	stock	file:/C:/Users/Mimi/Documents/Visual Studio 2010/Projects/ETFCE/ETFCE/Files/Product settings.pdf	28

Figure 24

Another important aspect of the information in the table is that there are the URL addresses of the two files added to the system's repository with some terms found there. This illustrates two main points:
all the files are accessible from the system despite of their different format,
and all the files contain some of the terms participating in the taxonomies.

Having annotations for each appearance of the terms from the taxonomy is the necessary basis for creating indexes.

5. Indexes

The indexing process consists of matching the added annotations back with the taxonomy elements and creating pointing indexes, connection between the annotations in text with the elements in the classification. As mentioned above for testing purposes there is a set rule that a term is considered relevant only if there are at least three of its connected terms in the same text. Then, an index in the system is added only for those terms.

Programatically this process can be implemented using several iterations, counting the number of child elements appearing in the list of annotations for each of the terms that have a parent role in the taxonomy. The only parent element in the sample is the term stock with IDTerm=3103. Looking at the results in **Figure 25** we see four indexes, all pointing to term 3103 and Document 19- Stock markets.txt. The reason to have indexes only for this document and this term is the number of annotations of the term's child items in the document. For Stock markets.txt there are eight annotations among which annotations for 3 of the child items, while for Product settings.pdf there are annotation only for the parent term.

IndexID	IDTerm	IDConstruction	Start	End	DocumentID
9052	3103	663	1	NodeImpl: id=3...	19
9053	3103	663	2	NodeImpl: id=5...	19
9054	3103	663	2	NodeImpl: id=5...	19
9055	3103	663	6	NodeImpl: id=1...	19

Figure 25

With tagging and indexing performed all the preparation is finished. All the added sources are annotated and all meaningful uses of the terms are indexed in the system.

6. Conclusion

This report covers a solution for intelligent tagging and indexing of unstructured data using a text engineering platform and the logical relations between terms in a taxonomic structure. It gives a slight idea how some ordered set of terms can be used to recognise realted appearances, but what is more important it shows the potential of an approach for further development. The more relations between terms we add and the more diverse meanings we imply in these relations, the richer the system's ability to interpret a text becomes.

References

1. M.Marzovanova, T.Tzoikova, Prof. D.Sc Valentin Kisimov,INFORMATION SYSTEM FOR TAGGING, INDEXING AND SEARCH OF UNSTRUCTURED DATA, 2012
2. M.Marzovanova, R.Georgieva,ПРИНЦИПИ ЗА ОРГАНИЗИРАНЕ НА ТАКСОНОМИЯ В ОБЛАСТТА НА ИКОНОМИКАТА НА БЪЛГАРСКИ ЕЗИК,2010
3. R. Rao, "From unstructured data to actionable intelligence," IT Professional, vol. 5, no. 6, pp. 29-35, 2003.
4. Robert Malone, "Structuring Unstructured Data," 2007.
5. Robert Blumberg and Shaku Atre, "The Problem with Unstructured Data," 2003.
6. Krishna Rupanagunta, David Zakkam, and Harsha Rao, "How to Mine Unstructured Data," Information Management , 2003.
7. Ashish Sureka, "Making Unstructured Data Findable Using Tagging and Annotation," 2006.
8. Hamish Cunningham, Yorick Wilks, and Robert J. Gaizauskas, "GATE - a General Architecture for Text Engineering,".
9. H. Cunningham, D. Maynard, K. Bontcheva, and etc. (2012, September) Developing Language Processing Components with GATE. [Online]. http://gate.ac.uk/sale/tao/split.html

Significance of Branding for Increasing Tourist Destination Attractiveness

Alexander Zyma[1], Yulia Lola[1], Sergii Kavun[2]

Kharkiv National University of Economics, 61166, pr. Lenina 9a, Kharkiv, Ukraine
zima@ksue.edu.ua, yulia-8-11-8@rambler.ru

[2] Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, 61174, pr. Peremogy 55, Kharkiv, Ukraine
kavserg@gmail.com

Abstract. The article considers the importance of tourism branding of the country, proved the relationship between the country's tourist brand as an element of the customer relationship management, companies and travel services, analyzed the level of brand development of the world as a tourist destination. Also in this manuscript are shown different statistical data for the tourist sphere for various countries as for the whole world, so and for Europe. Authors was developed the foundation of tourist brand in the form of a scheme of the relationship of different elements.

Keywords. Branding, country brand, tourist destination, tourist arrivals, Country Brand Index.

1. Introduction

Country brand strength is a nation's ultimate intangible asset and goes beyond its geographic size, financial performance or levels of awareness. Managed properly across every measure, it can be a lasting vehicle for goodwill, encouraging forgiveness in difficult times and disproportionately boosting the value of exports, from people to products to entire corporations. Therefore, a comprehensive study of country branding is important.

2. Literature Review

Many of scientists have studied the problem of the regions branding and rebranding such as Briggs S. (Briggs, 2001), Aaker David A. (Aaker, 1994), Durovich A. (Durovich, 2008), Kapferer Jean-Noel (Kapferer, 1994), Neumeier Marty (Neumeier, 2005), Lane Kevin Keller (Lane, 2007), Baker Bill (Baker, 2007), Kolb Bonita M. (Kolb, 2006), Pike Steven (Pike, 2008) and other.

Purpose of the article is to research for the influence of branding on the attractiveness of the destination. Arguably, a strong country brand is a driver of brand strength in other contexts – when a product, service or corporation is identified with a strong country brand, it has a better chance of premium pricing, longevity and preference in emerging markets. A weak country brand, like a weak product brand, leads to poor differentiation, ambiguous meaning and low recall in the minds of people who travel, invest and do business outside their borders. All of this affects a nation's ability to stand out regionally and globally and to realize future ambitions beyond its natural resources (UNWTO).

3. Statistical Data

However, some destinations do not have a brand strategy, and are supported by inconsistent advertising campaigns, creating a confused image to prospective customers. A strong brand is differentiated from others, has several strong advantages when compared to others, and has an attractive appeal to consumers. Moreover, there is a connection between the country tourist brands, region brand, brand of tourist companies and tourist services. Foundation of tourist brand is shown in Figure 1.

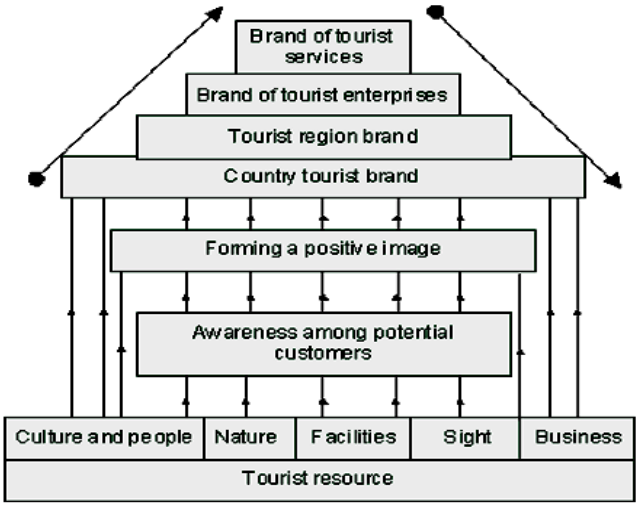


Fig. 1. Foundation of tourist brand (authoring development)

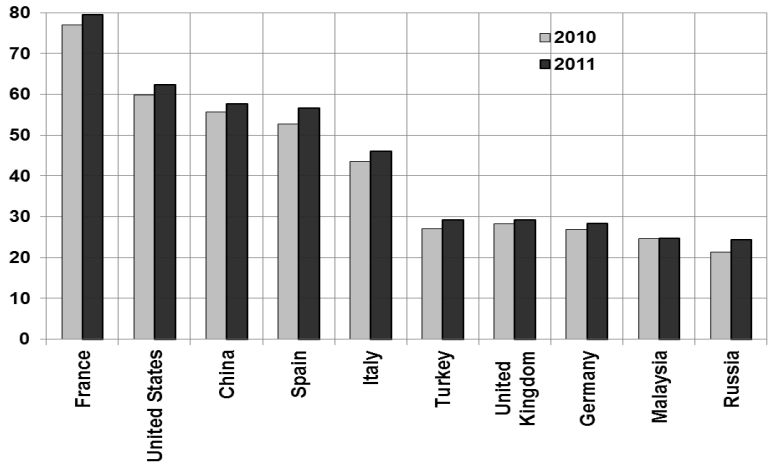


Fig. 2. International tourist arrivals for regions, in millions (Gadling site, 2012)

Tourist country brand is based on advertising, PR and promotion, must informal information obtained through word of mouth and personal recommendations. Tourist information centers of some country regions often do this (Lola et al., 2011).

Country branding as specific tourist destination is a process of forming a country image has a significant influence on increasing tourist flows and as a result on increasing tourist income to economy of the country. In 2012, there were 1 035 million international tourist arrivals, with a growth of 8,72 % as compared to 2010. South-East Asia (+9 %) were the fastest-growing regions in terms of tourist arrivals in 2012, followed by Central and Eastern Europe (8 %) (UNWTO, 2012).

International tourist arrivals for different countries are shown in Figure 2. In this figure are shown the indicators in percentages from total quantity of tourists for each researched countries in the world.

The first places on tourist arrivals occupied France, United States, China, Spain, Italy, Turkey and Germany that indicates the presence of strong tourist brands in those countries. Brands of countries are associated with a positive image of tourism, which is connected with a developed infrastructure and strong brand of tourist companies, which depend not only on the advertising, but also on their general management (Lola et al., 2010). Brands and touristic images of the different countries are presented in Table 1.

Table 1. Brands and tourist image of countries (according to the site adme.ru)

Country	Brand	Tourist image	Tourist company with strong brand
France	"See you in France"	Romance of Paris (about which we have heard from our youth), to visit the royal palaces (known from literature), see the Eiffel Tower, taste exquisite French cuisine	Accor Group is a French hotel group, part of the CAC 40 index, which operates in over 91 countries
United States	DiscoverAmerica.com "The United States welcomes all!"	Country of the democratic, huge skyscrapers, luxury hotels, unusual museums and fabulous entertainment	Marriott International Inc, Global Hyatt Corporation Starwood Hotels & Resorts Worldwide Inc, Hilton Hotels Corporation, InterContinental Hotels, Best Western Inc, Four Seasons Hotels & Resorts, Disney

China	"Wonderful China. Journey to the past and the future. China. Roads which remember".	Tourists are attracted by ancient traditions, the coexistence of socialism and capitalism in society, the great international exhibitions, the presence of 41 world heritage sites by UNESCO	China Peace International Tourism Co. Ltd. Marco Polo International Travel Service
Spain	 "Smile, you're in Spain"	Famous bullfighting and flamenco, a lot of museums, art galleries, palaces and other attractions, wonderful beaches	Barcelo Hotels & Resorts, Melia Hotels, Iberostar
Italy	 "Italy leaves its mark"	Wonderful beaches, a lot of museums, churches, palaces and other attractions	Domina Vacanze
Turkey	 "Turkey, Open for everyone"	Good service for the best price, excellent hotel animation, a lot of historical sites	Papillon Hotels, World of Wonders, Dolphin
United Kingdom	 "You're invited"	A lot of castles, churches, abbeys, museums, other historic places such as Stonehenge; famous queen family	British Airways, Thomas Cook Group
Ukraine		Through football, atomic energy of the USSR and the «orange» winter of 2004, Ukraine has created a fairly recognizable image	Reikartz Hotel Group

Ukraine is actively engaged the country branding as a tourist destination before the carrying of the international championship of Europe "Euro-2012". However, there was observed incoordination of the brand creation policy, as for example, were shown several

country logos that does not contribute to their memorability.

In addition, advertising campaigns, the creation of an image should be directed at a specific segment of the international tourist market. International tourist arrivals for Europe are shown in Figure 3.

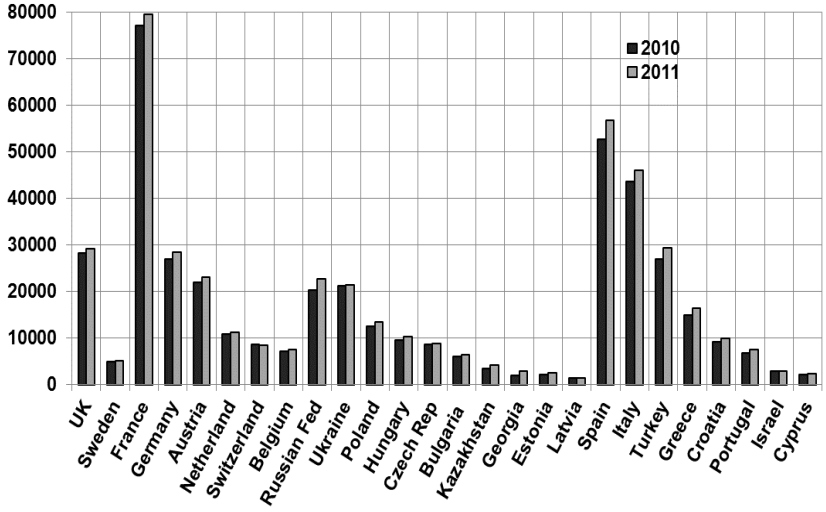


Fig. 3. International tourist arrivals for Europe (according to the UNWTO)

In this figure are shown the number tourists (in millions) from total quantity of tourists for each researched countries for Europe. As you can see in Fig. 3 Eastern European countries have relatively low tourist arrivals and weak brands. Ukraine is ranked second in terms of tourist arrivals (23 million) in this region of Europe. However, it is ahead of Poland, Hungary, Czech Republic and Bulgaria.

4. Results of Research

The vast majority of destinations worldwide posted positive figures, sufficient to offset recent losses or bring them close to this target. The recovery in international tourism is good news, especially for those developing countries that rely on the sector for much-needed revenue and jobs. International tourism demand held up well in 2010, despite persistent economic uncertainty in some major markets, the natural disasters suffered in some countries, political and social unrest in others, the serious disruption of air travel following a volcanic eruption in Iceland last April and the problematic weather conditions in parts of Europe and the USA in December. Tourism has once again proven to be a highly resilient sector.

Nevertheless, it is necessary to work closer and better towards increased integration and cooperation between all players involved in the tourism value chain to increase competitiveness and respond more effectively to challenges. 2010 also saw the rise in importance of mega-events – sport, culture and exhibitions – in terms of their extraordinary ability to attract visitors and position host countries as attractive tourism destinations. Notable examples include the Winter Olympics in Canada, the Shanghai Expo in China, the FIFA World Cup in South Africa, the Commonwealth Games in India, European Football Championship 2012 in Poland and Ukraine (LHT, 2012). These events lead to the increase in

the number of foreign visitors as in a certain moment and in the long time. Consider the example of Ukraine. Dynamics of international tourist arrivals for Ukraine are shown in Figure 4.

During international events in countries with weak tourist brand, the number of guests can be less than expected. So according to the dynamics of the foreign tourist arrival in Ukraine the function of trend line will be the following:

$$y = -21449x^3 + 1E+08x^2 - 3E+11x + 2E+14,$$

with a value of the reliability of the approximation $R^2 = 0,9129$.

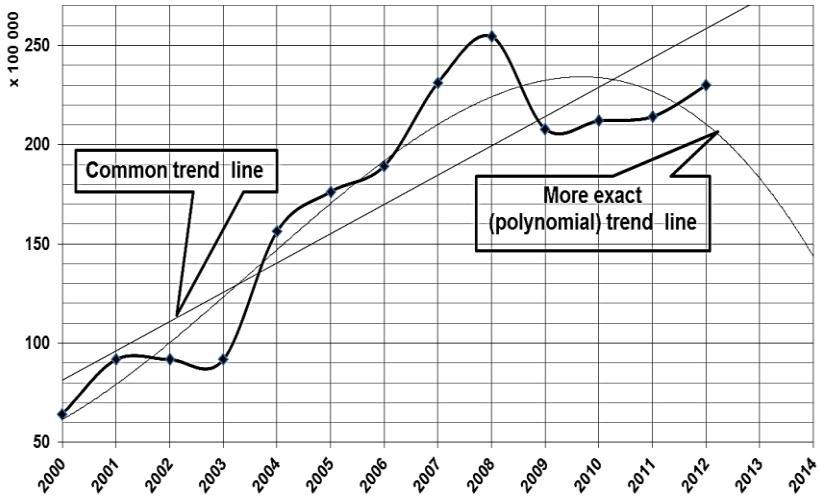


Fig. 4. Dynamics of international tourist arrivals for Ukraine (according to LHT)

The dependence can be described power function. By predicted about 26-30 million tourists can come in Ukraine for the entire 2012 (Zyma et al., 2010). However, championship results showed that this is unlikely to be possible. For such kind of event, a unique country brand was created or existing brand was reinforced. It can help to identify country among other and extract the strengths, advantages and attractiveness for travel.

To have the strongest country brand, it is necessary for country to be democratic, progressive, relatively politically and economically stable, and doing business in English, also the value systems and the freedom of communications: a major factor in world perception of a country and its culture, people, businesses and brands are important. The economic crisis is also a powerful factor in country brand strength, but mainly for those that avoided it. Countries, which managed to escape the worst of the banking collapse and maintain relatively strong economies throughout 2010, are Australia, New Zealand and Canada, which countries brands were defined as top three brands in the world in accordance with Country Brand Index (CBI, 2010).

A significant influence on a strong country brand forming has brand of some product or company, which headquarter is situated there. Despite crisis, economic downturn, political upheaval and public relations problems, countries with a clear identity, consistent values and a lasting heritage continue to perform at the highest level. For example, according to Country Brand Index brand China's position has actually dropped in the 2010 ranking while

simultaneously displacing Japan as the world's second-largest economy, showing that economic strength alone is not enough to build a strong brand.

Japan, on the other hand, has raised one position in the ranking, amid high-profile national PR problems such as the Toyota recall, economic slowdown and a weakened yen. Similarly, the BP crisis – widely feared to impact brand UK – seems not to have significantly damaged perceptions of the brand. Its continued high performance in heritage and culture as well as consideration and visitation show that legacy value and a strong identity as a nation can overcome austerity drives and corporate disasters. Similarly, while France has fallen this year, its biggest strength remains in Heritage and Culture, and it makes great use of that in exporting some of the world's most famous brands – from Chanel to LVMH Group – all of which continually reinforce its authentic history, culture and style. It is no coincidence, for example, that the emerging super-rich in China choose French luxury brands over others as emblems of their newly found economic and social status.

French brands such as Carrefour, Danone, L'Oreal, Evian and Renault also carry great appeal to the mass market, which reinforces a more populist aspect to French culture and specialization in retail, cosmetics and dairy products. France also continues to be a very strong tourist destination for the same reasons, remaining at number one in awareness and decision/visitation, and jumping to first place in familiarity this year. Undoubtedly, French music and cinema continue to influence global culture for the Francophone world, and France 24 and Canal+ are major vehicles for French culture (CBI, 2010).

5. Discussion

Like company, product or service brands, country brands create strong and positive impressions that generate desire and demand. The consistency of impressions and messages across media and channels keeps them top of mind and promotes them across audiences and categories of relevance, from business to travel and tourism. The leading country brands should have a healthy mixture of public and commercial broadcast networks with multiple stations, some international reach and a relatively free press. They also should have excellent communications infrastructure with high levels of Internet and mobile phone penetration. In a world defined by user-generated content, borderless communication through social networks, and unprecedented access to news, information and rich media, a country brand is now partly built by aggregated sentiments and content arising from people's personal experience. The more open a country, and the better its technological infrastructure, the more likely good experiences will permeate and impact awareness and preference for its brand.

Every country has specific landmarks, which are their city brand and thanks to means of communication make it famous for tourist. Some of these countries were included in the list of the top countries for 2012 by experts from Lonely Planet international journal (Lonely Planet, 2012) that presented in Table 2. For listed cities, new tourist destinations were developed or they were disclosed from other side.

It should be noted, that some of the city within one country could also have a distinct brand. It is associated with an outstanding or a tragic event, an attractive tourist object or natural and climatic characteristics.

The country's participation in travel ratings, joining the international travel and tourism association, including the tourism object in the List of World Heritage Sites (there are seven World Heritage Sites objects in Ukraine), influences the formation of the positive tourism image.

Country image as complex of emotional and rational conceptions, which based on specific country identification, own experience and private information, its social, political,

economic and financial standing, a level and kind of tourism development can form a strong country perception as tourism destination. Such kind of criteria allowed the defining the top 10 best value destination for 2011 which can cost a little presented in Table 3. At the ninth place of the best destinations, which are cost a little, is Ukraine. A good budget options it is possible to find in Lvov and Odessa are available for less than $50. Public transport is fantastic value, with a train between Kyiv and the Lvov costing less than $10, and a tram ride from the station into town a hundredth of that. Food and beverage is very cheap (Gadling site, 2010).

Table 2. The most attractive tourist country for 2011-2012 by experts from Lonely Planet international journal

The best country 2011	Description	The best country 2012	Description
Albania	Not so long ago, when the Balkans were considered an 'only for the brave' travel destination, only the bravest of the brave trickled into Albania. It has good cuisine, heritage sites, nightlife, affordable adventures	Uganda	This is the source of the river Nile – that mythical place explorers sought since Roman times. It is also where savannah meets the vast lakes of East Africa, and where snow-capped mountains bear down on sprawling jungles.
Brazil	Famous for samba, football and cinematic scenery, Brazil has always been known for celebration (Carnaval being the most obvious manifestation of this national joie de vivre). Yet, Brazil rings in 2011 with even more cause for jubilation. There will be the 2014 FIFA World Cup and the 2016 Summer Olympics in Rio de Janeiro,	Myanmar (Burma)	In 2010, the National League for Democracy revised its boycott to encourage independent travel (as opposed to package tours). As a result, Myanmar is set to be a hot new destination for independent travelers. Mountains and white-sand beaches, the kite-shaped country's most accessible area is the centre, which is filled with timeless towns and countless pagodas.
Cape Verde	Soaring mountains terraced in greens, a volcano with its head in the clouds, excellent watersports and sizzling, saucy festivals – but it was the sun that clinched the deal. With almost more days of sunshine than there are days in the year and with soft sandy beaches	Ukraine	Ukraine is co-host Euro 2012 and the four match venues have been selected to encourage further travel by visiting football fans. So Lvov becomes the jumping-off point for Carpathian exploration, while Kiev become base for forays to the Black Sea coast and the grim tourist attraction that is Chernobyl
Panama	For travelers, there is a return to the authentic – local heartland festivals, jungle treks and lodgings in sand-floor huts in the independent Comarca de Kuna Yala. With plenty of the country, still pristine, true adventure is only a boat or bus fare away. In 2011, Panama City gets greener, with the anticipated unveiling of the BíoMuseo.	Jordan	Petra, the ancient 'Red Rose city', Wadi Rum, Jerash and Madaba are adding weight to the country's tourism boom. One of the most open, friendly and welcoming nations in the Middle East, Jordan is an example to other states in the region of how to modernize while preserving cherished ancient traditions.
Bulgaria	Bulgaria has snow-capped peaks and great-value ski fields, golden-sand beaches and ancient port towns in Black Sea. Now it's ski slopes are de facto destinations for Europeans looking for cheaper alternatives, empty patches of lovely Black Sea beaches can still be found, and its quietly brilliant wine industry is flourishing	Denmark	Every year, Denmark tops a 'quality of life' list and is revealed as the happiest best-looking place on earth. All you have to do is hop on your bike – literally. Countrywide, you'll find around 10,000 km of bicycle routes. You can see the picturesque countryside or an architectural delight
Vanuatu	Vanuatu includes 83 islands. From mighty mountains and thunderous waterfalls to remote villages, from huge lagoons to	Bhutan	Buddhist Bhutan has always coyly shielded its charms from the wider world, but new areas of this remarkable

	tropical islets, there is so much on offer, far from the crowds. Don't expect ritzy resorts and Cancun-style nightlife; it's tailor-made for ecotourists		mountainous land are finally opening for business. Royal Manas National Park, prowled by some of the planet`s last remaining tigers, has reopened
Italy	There are 44 UNESCO World Heritage Sites here, more than in any other country. That Italy is celebrating only 150 years as a country in 2011. It still feels like a collection of regions. Each region has a pronounced character and qualities worthy of the small nations	Cuba	Cuba`s socialist credentials are gradually crumbling in the face of international capitalism. This is bad news for fans of colonial charm. The beaches will still be pristine 10 years from now. However, the country might not be quite so distinctive. Go while the clock is still stopped at 1959

Table 3. The top 10 best value destination for 2011 (according to the State Statistics Service of Ukraine and Lonely Planet journal)

Destination	Elements of cost
Bangladesh	Meals a day can cost less than $1, hotel room - $10 per night, excursion to the National park to watch for tigers - $150
Nicaragua	typical a price for hotel room - $30, the biggest expenditure is tax for flight from country - $32, which is included in the ticket price
Washington, DC (USA)	it can be explained by great amount of unique places which are free of charge – Lincoln Memorial, National Air and Space Museum, Capitol, Library of Congress, excursion to White house and State Department
Paris	The French capital is never going to win any awards for cheapness, but here is a winning formula that anyone can afford. This title was given to the city in terms of its price for meal. It is possible to buy a freshly baked baguette, a fist-sized hunk of cheese and a bottle of wine for $13.5 here
Namibia	Expenditures for meal, accommodation and transport per day are $50. Admission to the Cape Cross Seal Reserve on the Skeleton Coast costs $3 per person
The Philippines	Meals cost $20 per day, excursion by ferry - $5 for each hour
Argentina	A standard hotel room in Buenos Aires or on Patagonia costs $46 per night, supper in restaurant - $30. Terrace tickets for Boca Juniors, River Plate and others cost from $3.60
Naples (Italy)	Italy is not cheap country, but Naples is one Italian city that prefers food that is fresh, simple and good value. It is possible to find accommodation with meal for $100 per night, here is cheap street meal. The ferry from Naples to the idyllic island of Capri, one of Italy's classic journeys costs $14

6. Conclusion

Thereby, creating a strong the tourist brand of the country must take into account the following features:

1) The set of specific characteristics, emotional and rational impressions, and national peculiarities, social, economic and political environment can form a strong country image and as a result, it can lead to promoting a city or country as tourist destination by means of such brand on the international scale oriented on specific target group.
2) The process of forming a strong country brand should contribute branding of cities, tourist centers and travel agencies to obtain a synergistic effect;

3) The brand of the country is determined not only by the quality of tourism products, but also by the quality of their promotion;
4) Brand power of the country as a tourist destination affects the result of cultural, sporting and business services activities;
5) International ratings raise the brand and image of the country (Zyma et al., 2010).

Further and current theoretical (Lola et al., 2011) and practical researches (Zyma et al., 2009-2010; Kavun et al., 2010) should be directed to the positive development of the tourist brand of Ukraine, projection of branding strategy based on the country's cultural and historical heritage and the current situation on the international tourist market. It is also important to solve the problem of multi-choice when developing the advertising company in the country as a tourist destination (Kavun, 2010).

References

1. Aaker, D. A. (1994). *Managing Brand Equity*, The Free Press: New York.
2. Baker, B. (2007). Destination Branding for Small Cities. The Essentials for Successful Place Branding. Second Edition. Creative Leap Books: Portland. Http://destinationbranding.com/book&ei=GnIbUKnNAozHsgaGpYF4&usg=AFQjCNGmoAVU0PlAZvyYX2qaXEW4Y9jKcQ. Accessed September 1, 2012, 10:15.
3. Kolb, B. M. (2006). Tourism Marketing for Cities and Towns: Using Branding and Events to Attract Tourists. Butterworth-Heinemann: Oxford. Http://booklens.com/bonita-m-kolb/tourism-marketing-for-cities-and-towns. Accessed September 1, 2012, 10:15.
4. Briggs, S. (2001). Successful Tourism Marketing: A Practical Handbook. Kogan Page Publishers.
5. Durovich, A. (2008). Advertising in Tourism: How to promote the countries. Http://www.adme.ru/kreativnyj-obzor/kak-reklamiruyut-strany-262855/. Accessed September 1, 2012, 10:15.
6. Gomon, D. (2011). Vacations on the cheap. *Today*, Vol. 1, pp. 7.
7. Kapferer, J.-N. (1994). The New Strategy Brand Management, The Press: New York. Http://books.google.com.ua/books/about/Strategic_Brand_Management.html?id=sc1I27U4uigC&redir_esc=y. Accessed September 1, 2012, 10:15.
8. Kavun, S. and Zyma, A. (2010). Development of tourism as an influential factor of economic security in the region: status and prospects (for example the Kharkiv region). *Visnik L'viv State University of Internal Affairs. Seria: Economical*, Vol. 2, pp. 120-129.
9. Lane, K. (2007). *Keller Strategic Brand Management: Building, Measuring, and Managing Brand Equity*. Peachpit Press.
10. Launch: Country Brand Index, CBI. (2010). Future Brand. Http://www.futurebrand.com/wp-content/uploads/2010/11/CBI_BBC_2010_execsummary.pdf. Accessed September 1, 2012, 10:15.
11. Lola, Y.Y. and Kis, O.P. (2011). Organizational structure of the tourist information center in Ukraine. *Geography of recreation and tourism: theoretical and applied issues*, Vol. 14, 269-274.
12. Lola, Y.Y. and Novikova, M.V. (2010). Peculiarities of tourist enterprises' management activities. *Geography of recreation and tourism: theoretical and applied issues*, Vol. 11, pp. 47-54.
13. Lonely Planet. (2012). Http://www.lonelyplanet.com/ukraine/travel-tips-and-articles/76856. Accessed September 1, 2012, 10:15.
14. Neumeier, M. (2005). *The Brand Gap*: Revised Edition. Peachpit Press.
15. Pike, S. (2008). *Destination Marketing: an Integrated Marketing Communication Approach*. Elsevier: Amsterdam.
16. Gadling's site. Ten budget-friendly Caribbean destinations. (2010). Http://www.gadling.com/2010/11/22/ten-budget-friendly-caribbean-destinations. Accessed September 1, 2012, 10:15.
17. State Statistics Service of Ukraine. Http://www.ukrstat.gov.ua. Accessed September 1, 2012, 10:15.
18. World Tourism Organization UNWTO. Https://en.wikipedia.org/wiki/World_Tourism_Organization. Accessed September 1, 2012, 10:15.
19. Zyma, A. and Kavun, S. (2010). Event marketing in the tourism business. *Economy and Law*

chronicle, Vol. 1, pp. 61-67.
20. Zyma, A., Kavun, S. and Zyma, O. (2009). Especially the functioning of the social sphere systems. *Economic: problems of theory and practice*, Vol. 3, No. 256, pp. 814-820.
21. Zyma, A. and Scheyko, O. (2010). Event tourism: carnivals, festivals, celebrations, exhibitions, parks. *Preprint*. KhNUE: Kharkiv.
22. Zyma, A., Zyma, O., and Kuzminchuk, N. (2010). Index of competitiveness in the tourism. *Business Inform*, Vol. 8, pp. 42-45.

The Models of the Balanced Regions Development Based on Fiscal Policy

Olexandr Trydid[1], Lidiya Guryanova[2], Sergii Kavun[3]

[1] Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, Kharkiv, Ukraine, khibs@khibs.edu.ua

[2] Department of economic cybernetics, Kharkiv National University of Economics, Kharkiv, Ukraine, g_lika@list.ru

[3] Department of Information Technologies, Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, Kharkiv, Ukraine, kavserg@gmail.com

Abstract. In this article was considered models complex of balanced regions development based on fiscal (or tax-budget) policy, which will be able to give a possible for conducting structural imbalances analysis of the territories (regions) development in perspective time period and parameters determination of this fiscal policy, which can providing their elimination or preventing. This policy can be performed on account of a management of transformation processes for tax legislation and tax relations, tax levying and redistribution between territories budgets (it also calls tax policy), or on account of a management of the tax expenses, subsidies, subventions, budget investment (it calls budget policy). Proposed approach to form of financial manage scenarios of Socio-Economic Regions Development (SERD) based on cyclicality conception of the economic systems evolution, anticipative manage principles, and can account some parameters changes of the investment transfers distribution, in particular, region developed fund. It can give a possible of the aftermath's estimation for the different variants realization of fiscal policy, to implement a selection of an optimal scenario, which providing decreasing of the interregional socio-economic differentiation at saving of positive trend for national economy development in generally.

Keywords. Region, balanced development, fiscal policy, scenario simulation, simulation modeling.

1. Introduction

Social oriented strategy of a development of the economic space determines the following global priorities for a state regional policy, as a provision of a unified social standards, a reduction of the interregional economic differentiation. Among some tools of a state regulation for territories development (for example, creating of the special economic zones, infrastructure development, a preparing of the territories for an industrial mastering, a stimulation for an attracting investments, etc.), as is showing a practice of the regional management, the most effective are the financial tools, such as a forming of some founds of a financial support of the regions, a budget financing of a development for separated economy sectors of regions, tax policy. It should be noted that an existed fiscal (or tax) policy, which is suggesting of some re-distribution of the financial resource, leads to an inevitable infringement of some interests of region-donors, to slow downing of tempos of their economic growth, to absence of some stimulus for less developed regions, to reduce of the subsidization level, to increase of their own competitiveness. Some economic imbalances are inducing some politic

imbalances, and, as a result, a growth of social intensity, forming of threats in the border regions, an absence of any motivation for interregional interaction, and a strengthening of the centrifugal tendencies. Current situation or event is leading to necessary for a correction of the regional policy, in particular, the financial policy. One from some directions for improving of the financial regional policy is including a development of the model basis for a state financial regulation of the territories development.

In the modern scientific and economic publications is considering some approaches to simulation of the regional policy [1-10]. In the publication [1] has considered such simulation methods of the budget regulation mechanisms on the different hierarchical levels, as the econometrics methods, casual and non-casual approaches, panel data. The researches, were presented in [2], are considering some questions of using of the space lagged models for testing of an available overflow effect and possibility for decreasing of some costs, which linking with a stimulation of the region development. In the publications [3-4] is considering some fundamentals of the system-dynamic models of the social-economic systems, the simulation models of the region as a bilayer structure: first layer is a background economy, second layer is the specific large subjects (also known as agent-oriented approach). Publications [5-7] is considering some questions of a development of the optimization methods for region development strategies, analysis of the interregional economic collaboration based on complex (or set) of optimization interregional and inter-branch models. In the works [8-10] was offered the complex (or set) models, which gives a possibility for determination of the state regulation priority for ability to live spheres of the regional systems and to improve an effectiveness of a target coordination based on the multidimensional analysis and adaptive filtration methods.

It should be noted that although enough big interest to development of the models for forming of an effective regional policy, some approaches are remain unstudied, which allow estimating of a consistency for the tax, budget policy, investment processes dynamic and its influence on the convergence processes of the regional development.

2. Conceptual Scheme of Scenarios Development for the Financial Management of the Socio-Economic Regions Development

This research aims to develop the scenario models of socio-economic development of the regions, which allow to analyze the structural imbalances in the perspective of territorial development period and to determine the direction of the regional financial adjustment policies aimed at their elimination or prevention. The scenario is interpreted as a sequence of states of socio-economic systems of the territories due to the implementation of various options in financing regional policy.

Proposed conceptual model of scenarios development for the financial management of the socio-economic regions development is including the following main stages: 1) forming of an inertial scenario and characteristics changing of the socio-economic territories development based on realization of fiscal policy; 2) dynamic analysis of the region developed imbalances; 3) developing and analysis of some alternative managed scenarios of the SERD. Below the content of the stages scheme is considered in details, as it shows in the Figure 1.

At the first stage the inertial scenario of changes in the characteristics of socio-economic development (SED) of the territories due to the implementation of the adopted fiscal policy is developed. Meeting the challenges of this phase is carried out using the model alignment imbalances using tax instruments [11-12] and the proposed simulation model of regional financial regulatory [13-14]. Model of financial regulation of the territorial development has two main components: the unit of resource allocation, the unit of socio-

economic characteristics of the region.

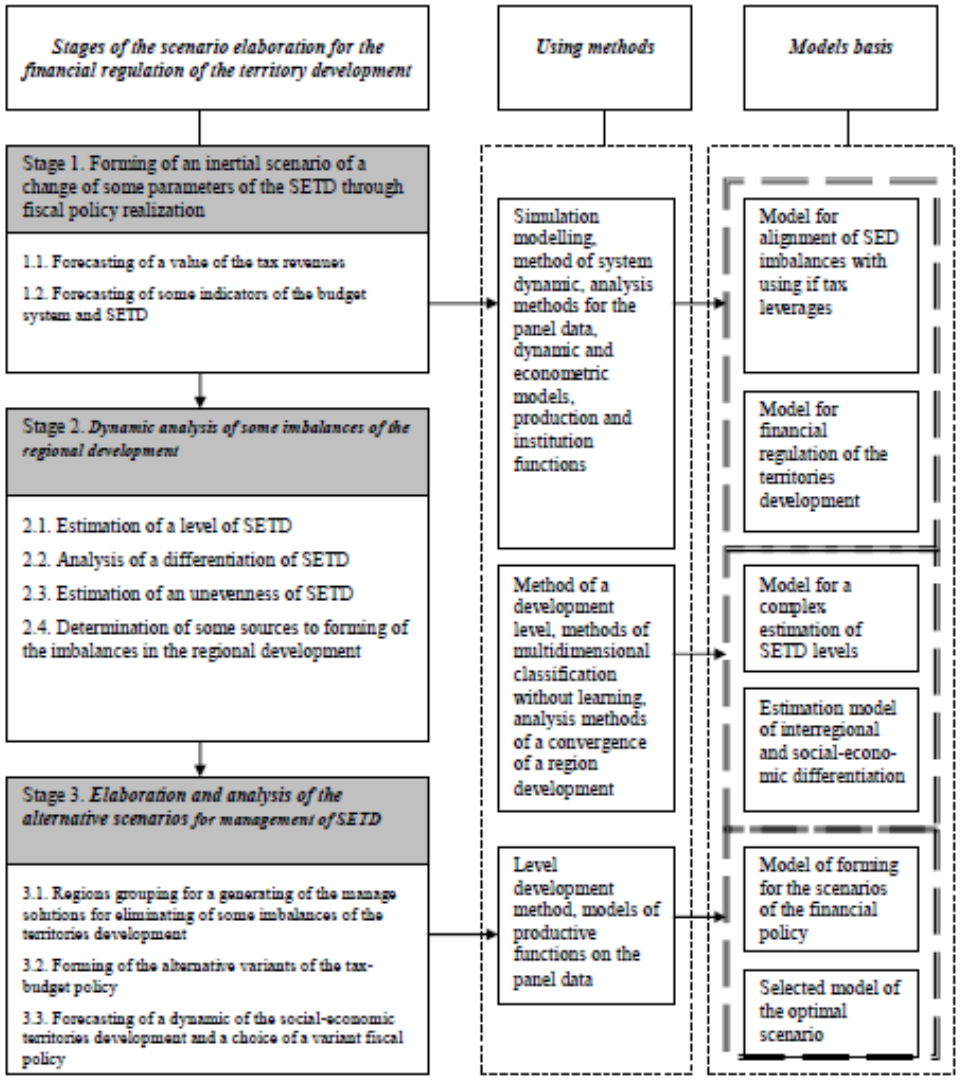


Fig. 1. The conceptual scheme of scenarios development for the financial management of the socio-economic regions development

The purpose of the first block is to simulate the possible value of the regional investment of transfers, subventions, grants. Targeting of the second block is to simulate the influence of the value of investment transfers, subventions, grants to the regions for the socio-economic development of regional systems. Therefore, simulation model of financial territorial regulation allows conducting multivariate projections of regional economic development and the state, depending on the adopted policy of the state of financial regulation.

Generalized model of state financial regulation of socio-economic development in the regions includes the simulation model of the resources allocation and 25 simulation models of socio-economic characteristics of the region. Outputs of this stage are the inertial scenarios of socio-economic development of the regions as a result of the adopted fiscal policy implementation. The pessimistic inertial scenario is to predict the parameters of fiscal policy with a time lag in the formation of positive "feedback" in the economy and, as a consequence, fiscal insufficiency. The optimistic inertial scenario accounts for the expansion of the tax base due to the changes in the parameters of the tax policy. Modeling of fiscal policy is based on the adopted under stabilization policy parameters of distribution grants, subsidies, investment transfers to the regions.

At the second stage, the analysis of the formation of imbalances in regional development is performed in the following areas: estimation of the regional socio-economic development, differentiation of the regional social-economic development, estimation of the imbalance of SED in the regions, identifying the sources of structural imbalances [15]. Assessment of the level of socio-economic development is conducted by using the method of construction the reference object - taxonomic indicator of development. The estimation of differentiation of SED is dynamic analysis of regional cluster formations, the analysis of individual propensity to migrate from regions with low levels of socio-economic development in to the group of regions with a high level of development. Evaluation is focused on the analysis of the regional structure of the upward or downward trend of the economic development. To estimate the irregularity the following data is used: coefficient of variation, coefficient of irregularity (differentiation), coefficient of imbalance, Tail index. Identifying the sources of structural imbalances is based on the decomposition of the Tail index. Herewith, the following factors of increasing the regional imbalances are considered: the unbalanced development of the groups of regions with high level of SED (donor regions) and regions with low SED (recipient regions), the unbalanced development of the regions with high level of socio-economic development, the unbalanced development of the regions of low level of socio-economic development.

At the third stage the alternative scenarios of managing the development of the territories are formed, aimed at eliminating or preventing the identified structural imbalances with saving the overall positive trajectory of the national economy development. The objectives of this phase are to generate a management decisions concerning the elimination of imbalances in regional development, the formation of alternative fiscal policy options, forecasting the dynamics of socio-economic development and selecting an options of fiscal policy. Solving the above tasks is performed by grouping the regions considering the following parameters: the level and rate of socio-economic development. Was allocated the following regions groups: regions-leaders, stagnant regions, developing regions and "problematic" regions.

Formation of alternative fiscal policy options suggests changing the parameters of the distribution of the investment transfers, in particular, the regional development fund, in between groups of regions. Since the cyclical downturn in the state investment policy is aimed at the increasing of the investment flows' speed, especially in the production of high added value, then adjustment of the parameters of the distribution of investment transfers is based on the research of the asset management ratio of the industrial and economic systems (PES) of the territories [16]. Prediction results of socio-economic development of the regions as a result of implementation of the different options of fiscal policy underlie the formation of alternative management scenarios of SED of the territories. The alternative compensational scenario assumes the estimation of aftermaths of the priority investment support of the regions – donors

with the implementation of pessimistic scenarios of tax revenues. The main target of the development of this scenario is to evaluate the possibility of forming a "compensatory" effect of reducing the depth of the economic crisis by changing the fiscal policy parameters. Alternative antirecessionary scenario is directed to model the results of the phased financial support of the recipient and donor regions. Financial support of the recipient regions allows reducing their level of subsidization and reducing the depth of the economic crisis at the beginning of the implementation of the state stabilization policy. The financial support of the donor regions is aimed at promoting an inward investment in the production of high added value and preventing the effect of "deferred" cyclical downturn in the forecasting period. Selecting an option of the fiscal policy is based on the analysis of the parameters of the regional financial policy that offers the alignment of the regional socio-economic development while maintaining the positive trend in the economy.

Thus, proposed above the conceptual scheme of scenarios development for the financial management of the socio-economic regions development can give a possibility for a consistency estimating of the tax, budget, and investment policy and its can improve a quality of an information and analytical base for management decision-making regarding of stabilization state policy.

3. Financial Management Scenarios of the Socio-Economic Regions Development

Developing the scenarios of socio-economic development of the territories is based on the model of financial regulation, which includes simulation models of budget system indicators and socio-economic characteristics of the regions. Simulation models are based on the dynamic econometric and panel data models. Panel data technique allows considering and analyzing of the regional differences in terms of the external "shocks" by effectiveness of the taken strategy. As a knowledge base for the modeling is used the Ukrainian budgetary system dated 1996-2010, socio-economic regions development dated 2000-2010. The simulation model of the indicators dynamic for the budget state system is including revenues of the consolidated budget, health care expenses, education expenses, social protection spending and social security, expenses for an economic activity, state budget expenses, subsidies, subventions, and investment regions transfers. In the simulation model of socio-economic characteristics of the development of territories are considered such variables as gross regional product, the total export volume, investment in fixed assets, the level of employment, the total import volume, the volume of innovative products, the value of foreign investments, the average monthly wage, income, level of economically active population, the provision of housing, bringing into service the apartments, the number of students at the universities.

Graphic of actual and forecast values of one from some indicators of the socio-economic regions development of Ukraine shows in the Figure 2. Similar results were also received for other indicators.

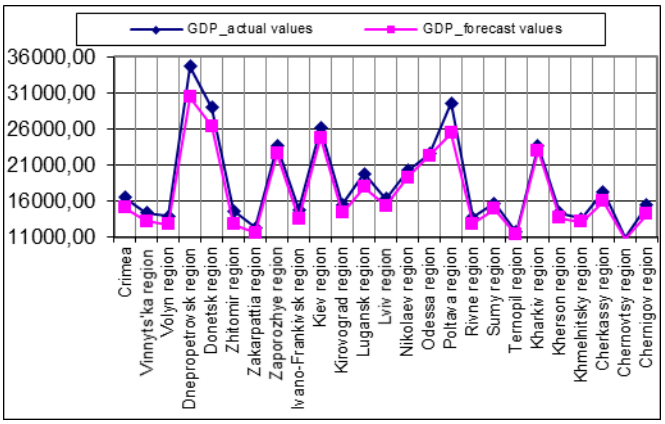


Fig. 2. Actual and forecast values of the GDP be personnel (UAH) by Ukraine's regions

An analysis, which was received based on simulation model results, is confirming quite high forecast precision and a possibility of their using for scenarios development of characteristics dynamic changes of the socio-economic territories development due to different variants realization of the tax-budget policy.

Optimistic inertial scenario (scenario 1) was developed, which takes into account the positive "feedback" in the economy, the growth in tax revenues by broadening the tax base due to changes in the parameters of tax policy and inertial pessimistic scenario (scenario 2), suggesting the formation of positive "response" in the economy with a time lag, the growth of the budget deficiency, which affects the financial capacity of leveling the socio-economic development of the territories, offsetting the negative effects of cyclical fluctuations through intergovernmental transfers.

The calculation results of the integral indicator of the level of socio-economic development of the regions, found based on the average values of the indicators of socio-economic development in the 25 Ukrainian regions in a forecast period and characterizing the tendency of the national economy development in general, are shown in fig. 3.

As can be seen on fig. 3 the adopted stabilization policy allows getting a positive effect on the medium term for two scenarios. This is evidenced by the dynamics of the integral index dated 2012-2014. The reduction of the integral index value in 2015 gives the opportunity to conclude that the downward trend and the emergence of the situation of "deferred" cyclical downturn is formed. The last confirms the necessity to adjust the parameters of the distribution of intergovernmental transfers in order to maintain the positive trends in the economy.

For forming of the alternative scenarios of the region policy was make grouping of the regions by a level and rate of the socio-economic region development. Region research by selected classification variables will allow the following groups: region-"leaders" (some regions with the high level and rate of SETD); stagnant regions (some regions with the high level and low rate of SETD); developing regions (some regions with the low level and high rate of SETD) and "problematic" regions (some regions with the low level and low rate of SETD). An analysis of the specific weight of an investment transfers, which was selected for these regions groups (as it shown in Fig. 4), allows us to make a conclusion that an accepted financial and region policy is directed, at least, to supporting of the stagnant and problematic regions.

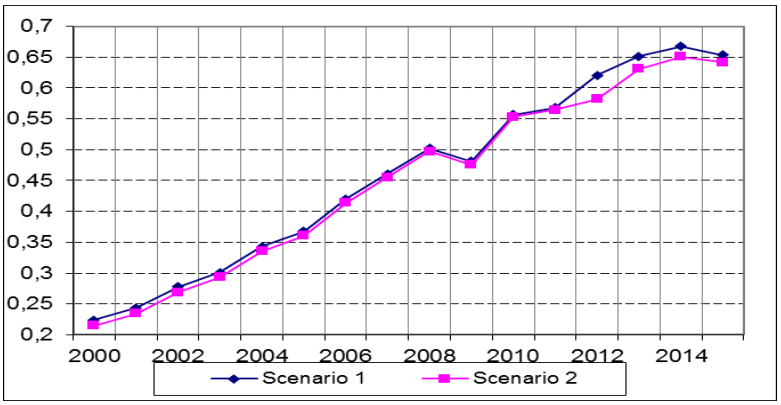


Fig. 3. The dynamics of the integral indicator of the level of socio-economic regions development

It should be noted, that an accepted policy for an levels alignment of the socio-economic region development, from one side, leads to a imbalances decrease in the region development on account of reducing of a level of a region depression, accelerated regions growth with low level of the socio-economic region development. However, from other side, it leads to a significant slowdown of a rates growth for region group with high level of the socio-economic region development. Because, as it was describe above, a priority of state investment policy under some conditions of cyclical downturn is increasing of the investment and innovation activity on the enterprises, which are making production with high value-added, then a parameters correction if the fiscal policy based on an analysis of the impact of resources in some region systems.

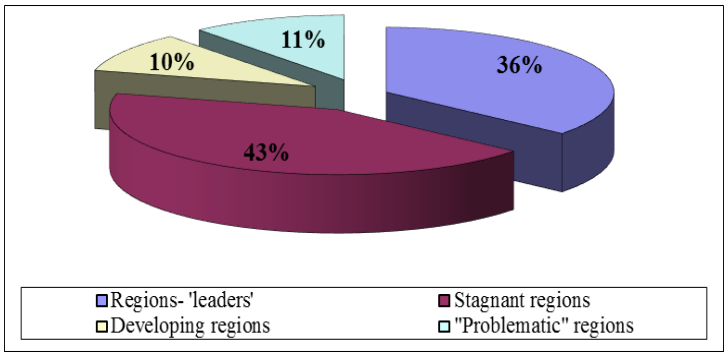


Fig. 4. Specific weight of an investment transfers, which was selected for the regions groups

As one from some alternative scenarios of the financial regional policy was considered the compensating scenario (scenario 3), which provides a stimulation of an economic growth not only for "problem" and stagnation regions, but and region-"leaders", which have a slowdown of the economic growth rates. At the same time, was considered a transformation possibility for the distribution mechanisms from 2013. As the original data for scenario forming also was considered a forecast of some tax revenue, which was received based on a model of the imbalances alignment of the socio-economic systems with using of the tax levers.

This analysis was conducted by pessimistic developed scenario; because this scenario is allowed to estimate a possibility about forming "compensating" effect based on a change of the budget policy parameters.

In the alternative anti-crisis scenario (scenario 4) was considered systematic financial supporting for the regions-recipients and regions-donors. At the simulation was accounted some investment transfers values in 2013, which were accepted some distribution parameters for the region-developed fund in the Budget codecs, which are oriented to the priority financial supporting for the "problem" territories. Correction of the distribution parameters of the investment transfers was carried out in 2014 with a purpose of a warning of a cyclic recession in a dynamic of the macroeconomic indicators (this recession is forecasting in 2015), which is directed to financial supporting not only for "problem" territories, but and for region - "leaders", for which are observed a significant recession of the economic growth rates.

Thus, these studies suggest the following conclusions: analysis of the predictive dynamics of the socio-economic development of the territories in case of the implementation of the optimistic scenario of tax revenue demonstrates the effectiveness of the adopted stabilization policy which helps to prevent the formation of a crisis in the dynamics of macroeconomic indicators and indicators of the regional development, to prolong the phase of growth up to 2014, to reduce the depth of the crisis in 2015; during the pessimistic scenario of the development of indicators of budget system "compensatory" effect of reducing the capacity of the financial regulation of the development of the territories can be formed by changing the parameters of the financial regional policy, which should be directed to support "problematic" regions as well as regions 'leaders', that are significantly slowing down the rate of economic growth during the "basic" policy of leveling the socio-economic development of the regions, the analysis of the coefficients of the unbalance of the socio-economic development shows a trend of convergence of the levels of economic development of the territories with different scenarios of development, reducing intergroup socio-economic differentiation.

4. Conclusion

Developed and proposed scenario models of the balanced territories development based on fiscal policy can give a possibility for a consistency estimation of the tax, budget, and investment policy and to increase a quality of the information and analytical base of management decision-making regarding financial stability policy, as for some regions, that for a state as a whole.

References

1. Lukyanenko, I.G., (2004). System simulation of the indicators of budgetary system of Ukraine: principles and instruments, PH "Kyiv-Mohylanska Academy".
2. Ponomarenko, V.S., Kizim, N.A., Klebanova T.S., (2009). Modern problems of modeling socio-economic systems: Monograph, PH "INZHEK".
3. Lychkina, N.N., (2011). Computer modeling of the socio-economic development in decision support systems. http://simulation.su/uploads/files/default/lych-comp-sim.pdf.
4. Fattahov, M.R., (2011). The agent-based model of socio-economic development of the megacities: PhD thesis, Central economic and mathematic institute, Moscow.
5. Ketova, K.V., (2008). Development of methods for the study and optimization of the development strategy of regional economic system: PhD thesis, Izhevsk State Technical University, Izhevsk.
6. Vasiliev, V.A., (2010). Edgeworth's equilibrium in a model of inter-regional economic relations, Siberian Journal of Industrial Mathematics, Vol. XIII, №1.

7. Vasiliev, V.A., (2009). About unlockable states multiregional economic systems, Siberian Journal of Industrial Mathematics, Vol. XII, №4.
8. Ponomarenko, V.S., (2004). The economic security of the region: analysis, evaluation, prediction: Monograph, Publishing House «INZHEK», Kharkiv.
9. Geyts, V., Kizim, M., Klebanova, T., Chernyak, O., (2004). Modeling of the economic security: state, region, enterprise: Monograph, Publishing House «INZHEK», Kharkiv.
10. Kizim, N., Klebanova, T., (2007). Adaptive methods in decision support: Monograph / Eds. Kizim N., Klebanova T., Publishing House «INZHEK», Kharkiv.
11. Yastrebova AS Modeling the process of alignment of the disparities in development of the regional systems using the tax levers / AS Yastrebova, O.V.Nikiforova, LA Chagovets / / Problems of the economy. - 2012. - № 2. - PP. 58 - 62.
12. Klebanova TS Modeling of the tax burden of enterprises in the transforming economy conditions: Monograph / TS Klebanova, Yastrebova AS - H.: PH "INZHEK", 2009. – 268 p.
13. Modern approaches to modeling the complex socio-economic systems / ed. VS Ponomarenko, TS Klebanova, NA Kizim. - H.: FLP Alexandrova KM, PH "INZHEK", 2011. – 280 p.
14. Models of estimation, analysis and forecasting of the socio-economic systems. Monograph / Ed. TS Klebanova, NA Kizim - H.: FLP Pavlenko AG, PH "INZHEK", 2010. – 280 p.
15. Models of estimation the unbalance and cyclical dynamics of the development of territories / ed. TS Klebanova, A. H Kizim. - H.: ID "INZHEK", 2011. – 352 p.
16. Klebanova TS Models of the differentiation of the competitive positions of the regions / TS Klebanova, L. Guryanova, EA Sergienko / / Competitiveness: problems of the science and practice: / Ed. d.e.n., prof. Ponomarenko VS d.e.n. prof. Kizim MO d.e.n., prof. O. Tishchenko - H.: FOP Liburkina LM, VD "INZHEK", 2009. – 264 p.

Analysis Models of the Interregional Socio-Economic Differentiation

Lidiya Guryanova[1], Tamara Klebanova[1], Robert Brumnik[2]

[1]Department of economic cybernetics, Kharkiv National University of Economics, Kharkiv, Ukraine, g_lika@list.ru, t_kleb@ukr.net

[2] Head of Center for Information Security, Institute for Corporate Security Studies, Ljubljana, Slovenia, robert.brumnik@ics-institut.si

Abstract. In this article was considered a model basis for an uneven estimation and cyclic dynamic of the socio-economic regions development (SERD), which are functioning at some considers of a high-level unevenness of an external environment, which renders a significant influence to increasing some risks and losses at a decision-making. This model basis is presented as a complex of the interrelated modules: a module of an estimation and analysis for unevenness SERD and a module of forming of the manage scenarios for SERD. Target orientation of the first module is an estimation for the interregional socio-economic differentiation, detection of some disproportions in a development of the separated regions. Content of the second module is forming of some manage scenarios for SERD, which are directed to decreasing of a level of interregional differentiation whilst providing a sustainable rates of economic growth. Also was formed a models complex of a differentiation of the territories development, which allow estimating of a stability of the cluster regions group by the SERD level, to make an analysis of their structural dynamic, to pick out a dominant kind of an asymmetry and priority structural components of the SERD for government regulation.

Keywords. Region, social and economic development, unevenness, structural disproportions, multidimensional analysis methods.

1. Introduction

The modern stage of development of the national economies, including Ukraine and the Russian Federation, is characterized by increasing of an unbalances economic space of any countries and by irregularity of the socio-economic development (SED) of regions. The unequal growth of individual elements that make up the socio-economic system of the region, leading to negative trends: does not give the opportunity to fully using a potential of an interregional cooperation, leads to increase a depth of cyclical crises, increasing the threat of disintegration. That the irregularity and cyclic development are considered as generally factors of destabilization, which reduces rates of an economic growth of regions.

Conditions for the functioning of regional systems are been characterized by a high level of uncertainty of the external environment that significantly affect to increase of risks and losses at making decisions. Therefore, these processes of SEDR management are poorly formalizable because they carried out in conditions of incompleteness, uncertainty and ambiguity of the original information. Since, building of systems of a regional development (SRD) should take into account the requirements of the dynamic stability, the reliability of functioning, the invariance with respect to external disturbances, the noise insensitivity to changes in the parameters, maneuverability, economy, etc. Addition of a factor of incomplete

information at the management of regional development is need to using of assessment models and analyzes strategies development on base simulation and forecasting of influence any different threats. The theoretical bases of development for models of the balanced SED are been reflected in many works of national and international authors [1-3, 10-12]. However, some questions for definition of a comprehensive quantitative assessment of SED level, assessment of an inhomogeneity for an economic space, identifying of source factors of the development asymmetry, construction of forecast scenarios for a development are not adequately reflecting in scientific researches.

2. Model Basis of an Uneven Estimation and Cyclicality Regions Development

Proposed models complex is including two main modules. Target orientation of the *first module* is an assessment of an interregional social and economic differentiation, to identify of imbalances in a development of individual regions. The content of the *second module* can be uses to forming some management scenarios of SEDR (areas), aimed to reducing regional disparity at a simultaneous ensuring of the sustainable rates of an economic growth.

The first module is including the following models:

Model of forming the indicators information space. The purpose of this model is to identify most important indicators of regional systems state. The initial indicators list was been formed based on data from the State Statistics Committee. It includes 52 indicators that to characterize the 14 structural components of the economic and social regions development, such as the "Industry", "Employment", "Agriculture" "Financial strength", "Housing", "Medical care", etc. Due to the not information provision many indicators during the analyzed period (2000-2009) was necessary to reduce the dimensionality of the indicators information space. We used a method "center of gravity", which gave the restrictions on the type of data and a sample size to build of such model. This method allows selecting the so-called signs represents for each structural component [7]. Using one of methods of factor analysis – a method of main components gave an opportunity to find the generalized latent factors that have the most significant impact on SEDR level [4, 13].

Model of grouping regions in terms of socio-economic development. The content of this model is an identification of homogeneous socio-economic characteristics of the regions groups for that can be developing differentiated variants for a regional policy. To build of this model can be used the cluster analysis methods, in particular, the iterative methods, which provide non-overlapping clusters, and which do not have restrictions on the number of objects and describing their symptoms [5, 7].

Model integrated assessment of the SED level. Target orientation of this model is quantitative (cumulative) assessment of SED region level. To build of this model can be used a taxonomic indicator of the level development, it is a synthetic value, which "resultant" from all indicators of SEDR [6, 7]. The choice of a method is due to its following advantages: the original system of indicators can include symptoms that have the different dimensions. In addition, the values integral index has a normalized variation range, which ensures the interpretability of these results. Comparative analysis of SEDR based on the obtained values of the integral index is conducting in both spatial, and space-time cuts. Integrated indicators provides some opportunities: to explore the structural mix of the identified groups of regions; to get a comprehensive assessment of the current and projected level of SEDR; to evaluate the possibility of strengthening of the inter-regional social and economic differentiation.

Model evaluation of irregularity and asymmetry of the development. The purpose of this model is an estimate of the heterogeneity of economic space. As the indicators of irregularity regional development are considering: the absolute and relative range of variation; the coefficient of the bundle; the quantile range of variation; the average linear deviation; the variance; the different coefficients of SED indicators. For analysis of the asymmetry of SEDR, as well as the intercompany imbalances are using the taxonomic analysis methods and a techniques of panel data analysis [7].

Model of cyclical development. On the one side, the current territorial structure differentiation leads to strengthen the instability dynamics of regional development and the appearance of cyclical ups and downs. However, on the other side, it leads to the polarization of the evolution of different groups of regions. To predict the cyclical dynamics of the region can be used the following methods and models: the models of the time series decomposition; the analytical smoothing trend; the approximation methods based on the Fourier analysis; the econometric methods (the regression analysis with panel effects and modeling production functions); the Markov processes (in particular, modeling transitions between the stages of the cycle); the methods of fractal mathematics and catastrophe theory; and the simulation. Predictive values of macroeconomic indicators to determine the potential of convergence and stability rates of socio-economic development of the regional groups [7].

The second module is including the following models:

Models of the forming variants of a financial regional policy. The purpose of these models is a determination of the possible volume of a public investment, or some subsidies, or some donations to the regions. For build a simulation model of distribute of some resources can be used a method of system dynamics by J. Forrester. An identification of some priority regional systems for public financial control of activity areas is performing on some base results of the analysis of the asymmetric development of the regions.

Model choice of strategies for a financial development of the regions. The content of this model is an estimate of the impact of different financial strategies for a social and economic dynamics of the region. Since the region is a complex socio-economic system, which includes the production, employment, finances subsystems, etc., then the models of the characteristics of these subsystems can be combined into a unified approach based on the system dynamics by J. Forrester [8-9].

Model assessment of the impact of the different variants of the financial policy. With the help of this model can build some scenarios of a change of some socio-economic characteristics of the regions through realization of the different variants of financial regulation: the priority financial support for regions with a high or low level of development; the uniform financial support for regions.

Model choice of the regional financial policy. The purpose of this model is to determine the variant of the financial policies. This variant should to reduce the inter-regional social and economic differentiation at the maximum rate of growth of GDP of the country in general.

Implementation of a set of models of SEDR provides an opportunity to balance the rates of growth of different groups of regions; also, it can reverse the adverse effects of foreign economic conditions on the dynamics of macroeconomic processes.

3. Regions Differentiations by the Socio-Economic Developed Level

One of main modules of this complex is a module of estimation of the interregional socio-economic differentiation. To achieve the tasks of the first module of the offered set,

which included: the grouping regions by SED level; the identifying of the numbered assessing and an analysis of the rate of growth of SEDR; an assessment of the irregularity regional development; an analysis of the asymmetry of SEDR when as inputs data are considered some space-time data of the social economic indicators from 25 regions of Ukraine and 79 regions of the Russian Federation for the period 2000-2009. With help of classification on base of the methods of cluster analysis was identified two groups of regions. The analysis of the average values for each group provides an opportunity to pick out the areas with high (H) and low (H) SED level. Optimality criteria's partition which addressing some group and intergroup values are confirm the correctness of that classification. Checking the quality classification of each study region to the selected group is accomplishing with help the discriminated analysis. Wilks's lambda values and the values found for posterior probabilities indicate the 100% correct classifications.

Based on the methods of taxonomy also was determined the values of an indicator of SEDR level for the period from 2000 to 2009 (on example of Ukraine). Positive growth on the indicator of SEDR's level are observed in 14 regions, it's representing above 56% (Figure 1).

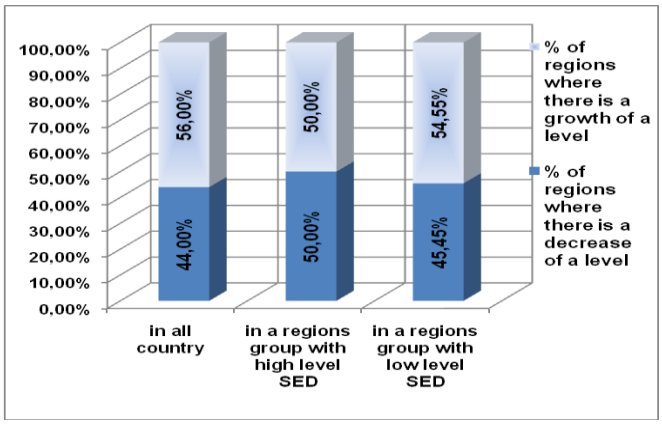


Fig. 1. Changes in the level of development of the regions of Ukraine for the period 2000-2009

Analysis of the share of regions with positive growth indicator values in the group with high level development is observed 50% of the total, and in the group with low level development – 45,45%. The leading regions with the highest values of the index level of SED are Kharkiv, Odessa, Dnepropetrovsk regions. Similar studies were been conducted in regions of the Russian Federation. The share of Russian regions with dynamics of the growth of the level of development are 22,08% of the total (Figure 2). The positive growth trend of this indicator in the group of regions with high level development is observed 16,67% of the regions, and in the group with a low level of development – 22,54% of the regions. Leaders by SED level from 2000 to 2009 are Tyumen, Moscow and Sakhalin regions.

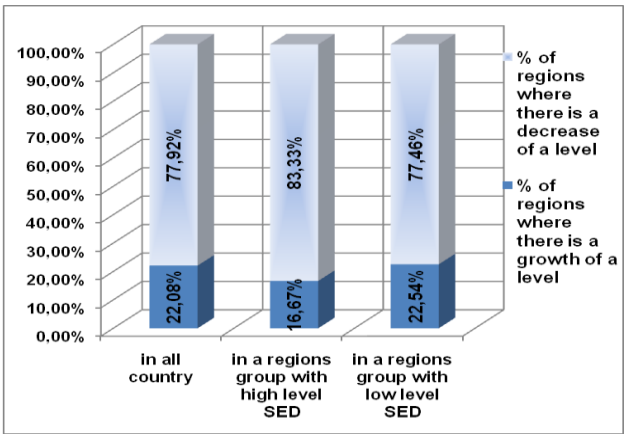


Fig. 2. Changes in the level of development of the Russian regions for the period 2000-2009

Average values indicate about higher SEDR level (on example of Ukraine) as compared with other regions (on example of Russia) during the analyzed period. Visualization of results related to the distribution of the regions (on example of Ukraine) according from the level of development and it is based on the cartogram (Fig. 3). On this cartogram are show rank region that determines its place in the general population, depending on the founded value for the level of development.

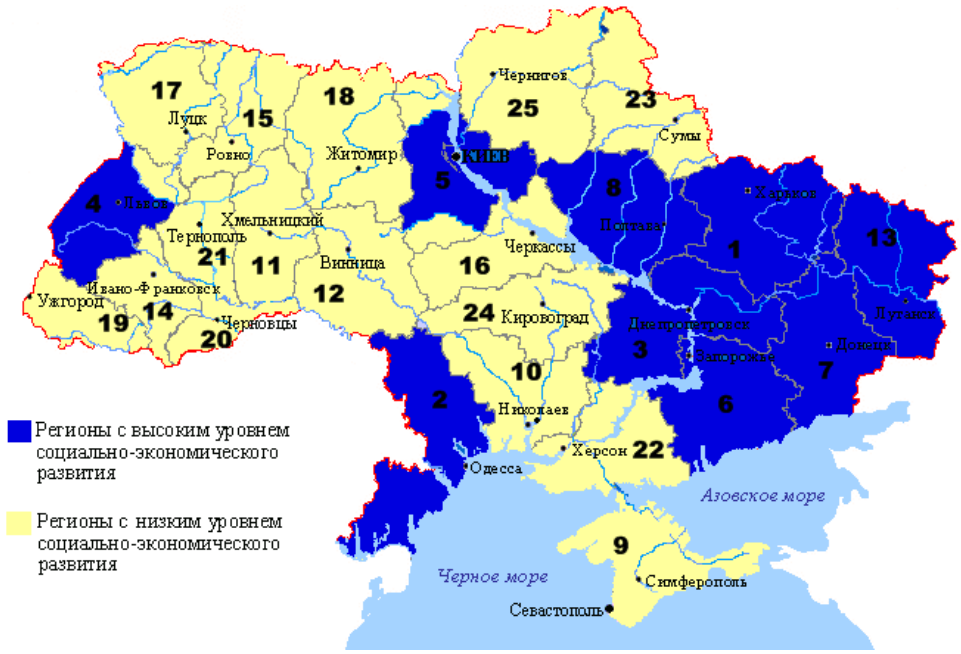
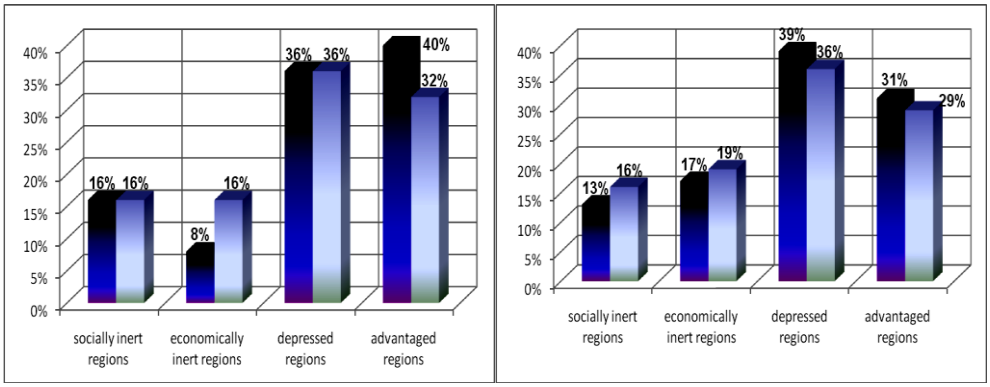


Fig. 3. The map of the ranks of SED
(blue color it's the regions with high SED level; yellow color – with low SED level)

Analysis of the growth of SED of Ukraine has shown that high growth rates have Kirovograd and Chernovtsy regions, in which the value of the index in 2009 compared to 2000 increased 2,89 and 2,81 times respectively. Such growth rates have led to a significant reduction of an inequality in the group of regions with a low level of development. In addition, those rates improved the balance of social and economic environment in the country as a whole. Within Russia, the highest growth rates of SED have Chukotka autonomous district.

A comparative analysis of SEDR makes a possibility to determine the qualitative features and directions of the development regional groups, "advantaged regions" – characterized by the values of the levels of SED above the average. For those regions, i.e. regions with high level of SED, the current task is a minimize costs to achieve a certain social result; "economically inert regions" – compared to the average it's a low the value of the level of economic development, but a high level of social development. The most important task of the group is to maintain a balance between economic opportunities and social needs; "socially inert regions" – compared to the average it's a low level of social development, but a relatively high level of economic development. This situation is due primarily to lack of attention to social issues; "depressed regions" are characterizing by values of the levels of SED below the regional average. These regions do not have equity any transfers for normal development. For these regions is characterizing by the full range of social and economic problems. Examples for Ukraine and Russia, throughout the study period in the group prosperous regions include the regions which shown in Fig. 4.



a) Ukraine b) Russia

Fig. 4. Region's share of each group in 2009 compared to 2000

Thus, the analysis of the economic space of Ukraine and Russia confirms the growing imbalance and asymmetry, as in SEDR, its can slow down the pace of SED of both countries.

4. Conclusion

The paper discusses the conceptual basis of the provisions and the model estimates of the irregularity and cyclical dynamics of the socio-economic development, the implementation of which is aimed to reducing an interregional disparity while ensuring sustainable economic growth; was proposed the complex of classification models of the territorial development which is based on some modern methods modeling and forecasting is giving an opportunity to

assess the stability of some cluster formation regions at SED level, to analyze their structural dynamics; to established a system of the integrated indicators of SEDR, allowing doing an analysis of trends in the regions development, to determine the asymmetry types of regional development for the different regions, highlight the dominant type of asymmetry and the priority structural components of SEDR for state regulation

References

1. Barro, R., Sala-i-Martin, X. (1991), "Convergence Across States and Regions", Brookings Papers on Economic Activity, Vol.1, pp. 107-182.
2. Cuadrado-Roura, J., Mancha-Navvaro, T., Garrido-Yserte, R. (2000), "Convergence and Regional Mobility in the European Union", 40th Congress of the European Regional Science, Barselona, 2000, available at: http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa00/pdf-ersa/pdf/450.pdf (accessed 23 September 2013).
3. Geyets, V., Kizim, N., Klebanova, T., Chernyak, O. and others (2006), Modeling of economic safety: state, territory, enterprise: Monograph, Publishing House «INZHEK», Kharkiv.
4. Kavun, S. (2007), Informational safety in business activity: Monograph, Scientific edition, Publishing house KhNEU, Kharkov.
5. Klebanova, T. and others (2007), Mechanisms and models of crisis management: Monograph, Publishing House «INZHEK», Kharkiv.
6. Klebanova, T., Guryanova, L., Sergienko, E. (2009), Models of the differentiation of the competitive positions of the regions. Competitiveness: problems of the science and practice, Publishing House «INZHEK», Kharkiv.
7. Klebanova, T., Kizim, N., Guryanova, L. and others (2011), Models of estimation the unbalance and cyclical dynamics of the development of territories, Publishing House «INZHEK», Kharkiv.
8. Klebanova, T., Yastrebova, A. (2009), Modeling of the tax burden of enterprises in the transforming economy conditions: Monograph, Publishing House «INZHEK», Kharkiv.
9. Klebanova, T., Guryanova L., Bogonicolos N., Kononov O., Bersutskiy Ya. (2006), Modeling the financial streams under conditions of uncertainty: Monograph, Publishing House «INZHEK», Kharkiv.
10. Lopez-Rodriguez, J. (2008), "Regional Convergence in the European Union: Results from a Panel Data Model", Economics Bulletin, Vol. 18, No. 2, pp. 1-7.
11. Lukyanenko, I. (2004), System simulation of the indicators of budgetary system of Ukraine: principles and instruments, Publishing House "Kyiv-Mohylanska Academy", Kyiv.
12. Lychkina, N. (2009), "Computer modeling of the socio-economic development in decision support systems", available at: http://simulation.su/uploads/files/default/lych-comp-sim.pdf (accessed 23 September 2013).
13. Ponomarenko, V., Klebanova, T., Chernova, N. (2004), The economic security of the region: analysis, evaluation, prediction: Monograph, Publishing House «INZHEK», Kharkiv.

Analysis of the Impact of Migrants' Money Transfers on Economic Development - Pro and Contra

Valentina Tolpinschi

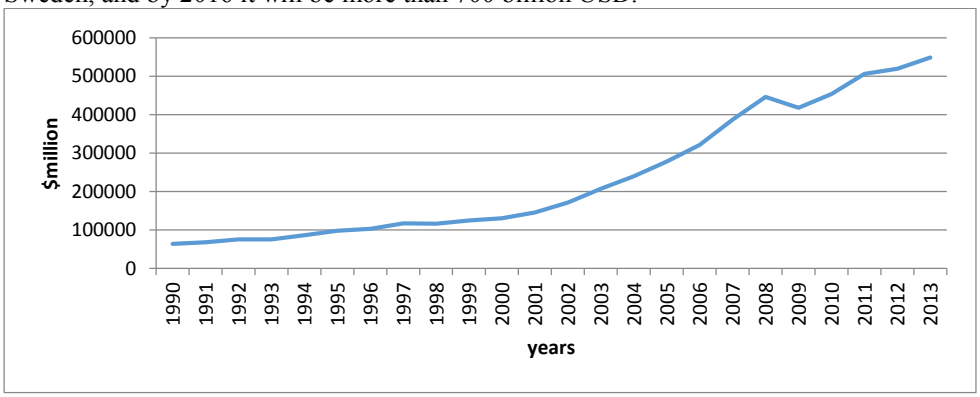
Academy of Economic Studies, Moldova

Abstract. It is often said that remittances are a much more stable source of external financial impact that it might seem at a first glance. Guest workers' international money transfers may noticeably affect budgets and financial systems of many countries. This money acts as a lifeline for the poor, increasing income for individuals and families. While migration can have both positive and negative economic, social, and cultural implications for countries of origin, remittances are the most tangible and least controversial link between migration and development. This research focuses on the study of the impact of remittances on the gross domestic product of the Republic of Moldova from 1999 to 2013.

Keywords: migration, Remittances, money transfer, impact, Economic Growth, state budget.

Migration and money transfer tendencies in the world

Money transfers are nothing new to the history. The most significant increase in guest workers quantity from 1990 till now was registered in the US. Their number has increased from 23.3 to 45.8 million people during that period. According to the World Bank, there are around 232 million guest workers today, which is roughly 3 percent of the current total world population. The amount of money transfers grows along with the number of migrants in 137 countries. According to Pew Research Center, in the later 13 years this amount tripled, and totaled 511 billion USD, which is comparable to GDP of some small countries like Poland and Sweden, and by 2016 it will be more than 700 billion USD.



Source: World Bank

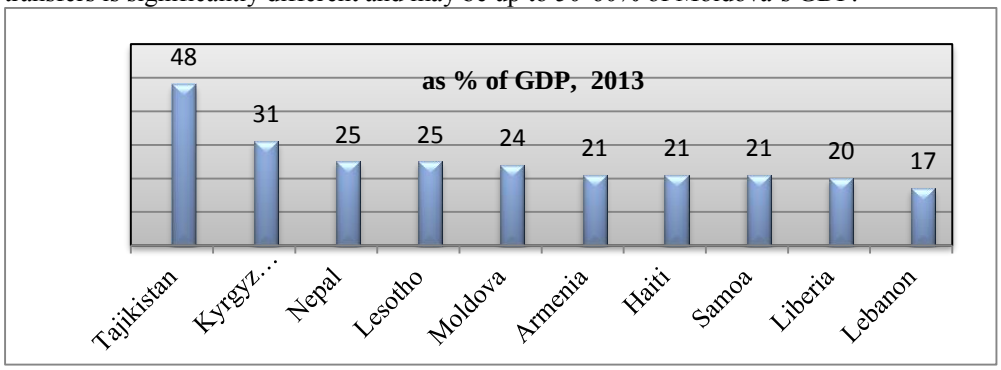
Figure 1. The dynamics of money transfers in 1990-2013 years

USA and the European Union remain the most popular immigration targets for the last decades. According to a research by European University Institute, there are up to 43 million migrants in the EU. 20 million are the Europeans who live in a country different from their country of birth, the rest are from outside of the EU. The EU is in a close second place after the absolute leader – the US.

It is not surprising that countries sending the most immigrants receive the most money transfers. The top recipients of officially recorded remittances for 2013 are India (with an estimated $71 billion), China ($60 billion), the Philippines ($26 billion), Mexico ($22 billion), Nigeria ($21 billion), and Egypt ($20 billion). Other large recipients include Pakistan, Bangladesh, Vietnam, and Ukraine.

Two thirds of money transfers go to developing countries. The most notable feature is the fact that all the declared money transfers are at least three times as large as the official development support in these countries. Some countries' (including Moldova) macroeconomic indices and general economic wellbeing heavily depend on immigrant money transfers.

The experts of the World Bank calculate the share of personal transfers in countries' GDP to demonstrate their economic dependency on this kind of monetary flows. Moldova is among the top five leaders having 24% during the last two years (the National Bank provides a slightly different number – 23%). Independent experts claim that the amount of illegal transfers is significantly different and may be up to 50-60% of Moldova's GDP.



Source: World Bank Remittance Prices Worldwide database

Figure 2. Remittances as percent GDP

For example, almost a half of Tajikistan's GDP is provided by guest workers in other countries.

The World Bank has created a database of the costs of sending and receiving small amounts of money from one country to another. It costs over 47 percent to send $200 from Tanzania to Kenya, the most expensive corridor, but only 4.6 percent to send the same amount from Singapore to Bangladesh, the least expensive one. If the cost of sending remittances could be reduced by 5 percentage points, remittance recipients would receive over $16 billion more each year.

Remittance costs are the key determinant of resource flows to developing countries, and lowering these costs is an important policy objective, as affirmed by the G20 in their 2008 commitment to reducing the global average remittance cost by 5 percentage points in 5 years.

Table 1.

Remittance prices

5 Most costly Corridors in USD	Average cost	5 Least costly Corridors in USD	Average cost
Ghana > Nigeria	27,12	Singapore > Bangladesh	1,41
Tanzania > Rwanda	19,97	United Arab Emirates > Pakistan	2,15
South Africa > Mozambique	19,06	Saudi Arabia > Pakistan	2,21
South Africa > Angola	18,67	United Arab Emirates > Sri Lanca	3,15

Source: World Bank Remittance Prices Worldwide database.

The global average total cost for sending remittances was 8.58 percent of the total amount sent in the first quarter of 2013, as measured by the World Bank's Remittance Prices Worldwide (RPW) database[5] . The cost of remitting from the G20 countries decreased in 4Q 2013, and now stands at 8.16 percent, down from 8.72 percent in the previous quarter. The average cost of sending money to the G20 countries that are included in RPW as receiving markets has fallen to a record low 8.86 percent in 4Q 2013 – a significant decline since the previous quarter when it was recorded at 10.57 percent and below 9 percent for the first time since the indicator was introduced in 2008.

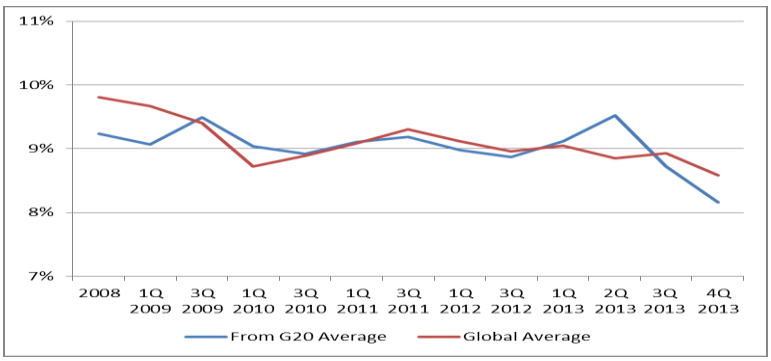


Figure 3. Average cost of sending USD 200 from G20 countries

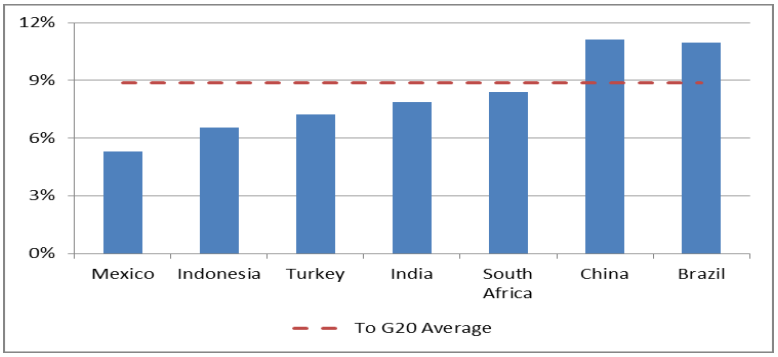


Figure 4. Average cost of remitting to G20 countries in 4Q 2013

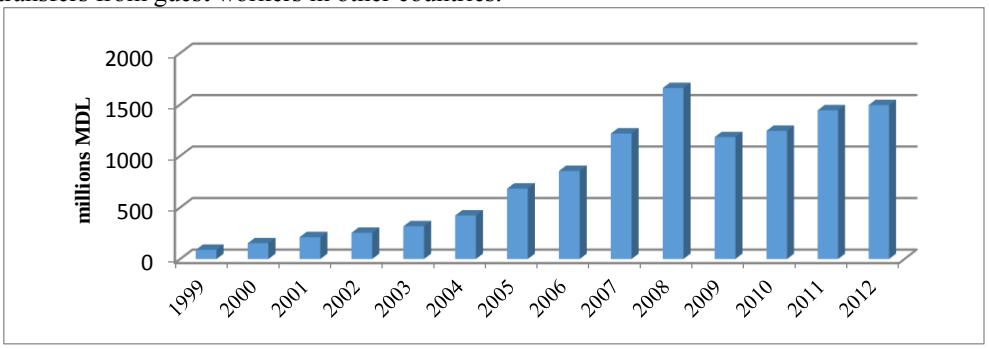
China and Brazil are the most expensive countries in the G20 to send money to, with average costs of 11.13 and 10.97 percent respectively (figure 4). Mexico and Indonesia remain

the cheapest receiving markets in the G20 group, with averages of 5.29 and 6.53 percent, respectively.

Strong growth continues in Russia, supported by high oil prices, underpinning buoyant remittances from migrants in Russia to Tajikistan and Ukraine. The majority of migrants from Romania, Russia and Serbia are, among others, in Western Europe, where economic growth has, however, been weak (GDP in the Euro Area contracted 0.5 percent in 2012 and unemployment has been rising). Remittances to Romania and Poland have gyrated in recent years. They surged after the countries were accepted into the EU in 2004 but dropped significantly after the crisis in 2008, partially due to increasing numbers of migrants returning home.

Remittances to Moldova

There are around 1 million Moldova citizens working abroad. Russia is seen in a favorable light compared to Western Europe by immigrants because due to stable oil price growth was able to support millions of guest workers from CIS. Transfers from Russia to Armenia, Georgia, Kirgizia, Moldova, and Tajikistan have grown in 2012. In contrast, some of the countries of Eastern Europe (e.g. Serbia, Albania, Kosovo, Romania) received less transfers from guest workers in other countries.



Source: National Bank of Moldova

Figure 5. Remittances to Moldova in 1999-2012 (US $ million).

According to the National Bank of Moldova, the amount of money transfers from abroad totaled 1494.23 million USD in 2012, which is 3.5% more than in 2011. It is important to note that 88.9% of them were made with international money transfer systems, and the rest used bank accounts.

At the same time, around 1.4% of country's GDP is transferred from Moldova to other countries. Our country is 38th by this criterion in the world. The largest part of outgoing monetary transfers is registered in Luxembourg (19.2% of GDP).

Salary constitutes 42.5% of natural persons' income, while money transfers from guest workers in other countries constitute 16.6%, still being a very substantial source of income. Villagers are much more dependent on money transfers, compared to people living in cities, having 21% of their income constituted this way, while for the cities it is 13%.

World Bank experts forecast that money transfers will decrease in 2013. It is expected that total transfers aren't likely to reach their historical maximum, registered in 2008, which constituted 1.66 billion USD.

The largest part of money transfers to Moldova is from Russia and Italy, but the change in their structure suggests fluctuations in European labor market.

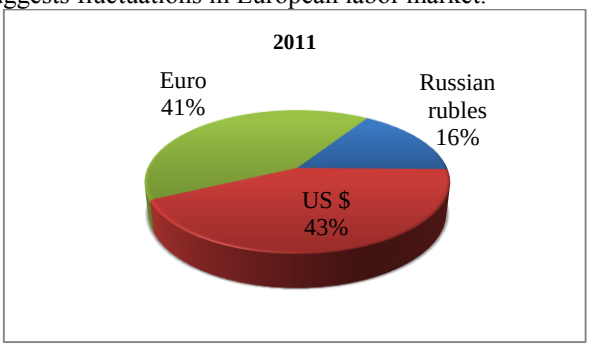


Figure 6. Remittance currencies in 2011

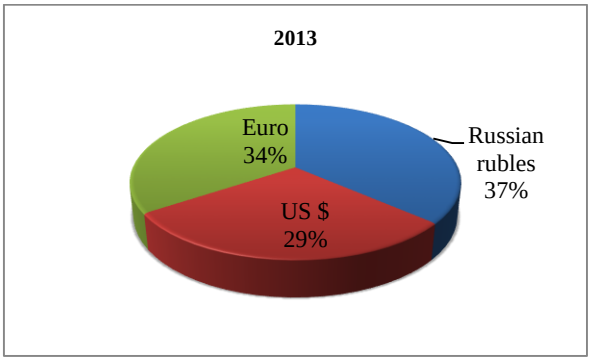


Figure 7. Remittance currencies in 2013

Thus, if in 2011 42.7% of transfers were in USD, 41.2% in euro, and 16.1% in Russian rubles, the situation changed in the third quarter of 2013: the majority of transfers were in rubles (36.4%), Euro being close second (34.3%), and USD seeing a substantial decrease (29.3%).

Now if we analyze geographical origins of money transfers from guest workers, Russia is in the lead with 65.6% (which is 4.1% more than in 2011), Italia – 9.1% (1.4% less than in 2011), Israel – 5.4%, the US – 4.9%, France – 1.3%. These five countries total 86.3% of all transfers.

According to the World Bank, 65.6% of money transfers come from Russia. Italia is the second with 9.1%; Israel is the third with 5.4%. Average transfer amount from Russia to Moldova was 525 USD (516 USD in the previous year). By the way, money transfers from Russia are one of cheapest in the world (around 2.4% of the transfer). According to the World Bank, world average cost of a transfer lowered from 9.3% in 2008 to 8.96 by the end of the third quarter of 2012.

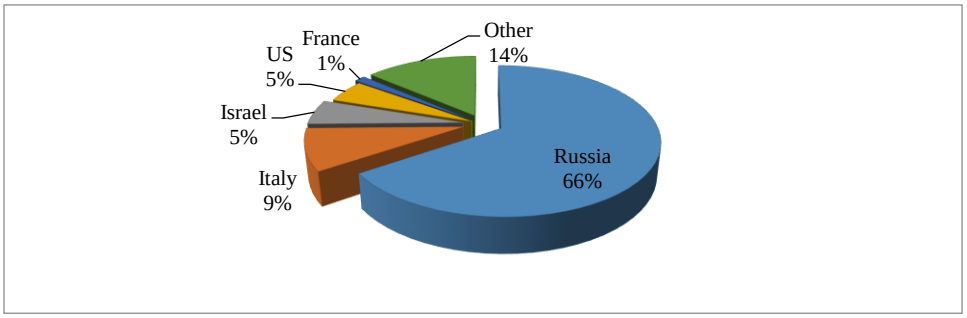
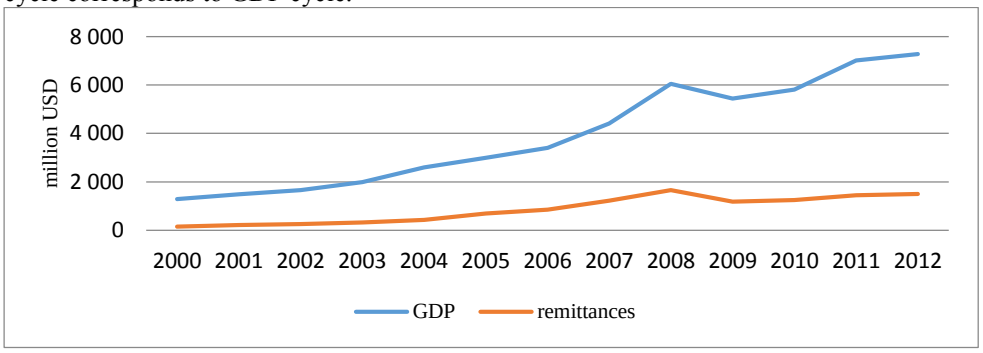


Figure 8. Money transfers origin countries in 2012.

Mass migration, which started at the end of the previous century, was conditioned by a difficult socio-economic situation in the country. At that time in seemed that it may have positive impact for the country: people will earn money, return to the country with a different mentality, and will develop their business and country's economics. Unfortunately, neither current, nor previous governments troubled to create favorable conditions for this. More and more migrants do not want to return. Furthermore, they do not want to invest in Moldova for the fear of increased risks, related to bureaucracy and corruption.

Monetary transfers and development in Moldova

Money transfers facilitate economic development of Moldova, money transfers activity cycle corresponds to GDP cycle.



Source: National Bank of Moldova

Figure 9. Corresponds remittences activity cycle to GDP cycle in Moldova (million USD).

Unfortunately, Moldova receives additional income in the form of money transfers and it is spent on everyday needs, long-term goods and real estate, not on business investment. Even the use money transfers for consumption stimulates economic development. Thus, a high level of goods and services consumption is maintained. But in this case the correlation between spending and savings and the residence of recipients (village or city) are of uttermost importance.

Impact of Remittances on Sustainable Economic Growth

The merit of remittance flows might lie more on increasing the level of income for the poor rather than the growth of the economy as a whole.[4] The economic role of money transfers is most of all evident in the countries that lost a substantial part of their population due to migration and have quite low average income. The role of transfers as an important and stable source of external financing of economic development was stressed in 2003 in the Global Development Finance[3] paper by IBRD and was acknowledged by regional development banks.

1. The data in Central Asia and Moldova demonstrates that monetary transfers may negatively influence child development. Prolonged lack of one of the parents (especially mother) exerts negative influence on study progress at school, social behavior and psychological wellbeing of kids remaining home.
2. The main issues faced by immigrants are related to legalization of prolonged stay at host countries and acquisition of work permit. Russia toughened the measures of migration control at the end of 2013: a limit on the share of guest workers in retail was set, monetary transfers by migrants to their homeland is to be taxed, etc. Moldavian diaspora reported that during only a couple of weeks around 3500 Moldovan citizens were deported. And that may have substantial negative impact on Moldova.
3. The migrants from European countries may encounter difficulties due to toughened hostile rhetoric and migration policy. Another problem might be high unemployment ratios in some of European countries. Italy, Spain, France, Greece, and Great Britain expect a higher unemployment growth rate among migrants than among native citizens. In Germany, Austria, Belgium, Canada and the US, on the contrary, migrants are promised higher employment.
4. Low quality of official data evidences the fact that labor migration and monetary transfers related to it, for the most part remain a black box for the political leadership on different levels. To consolidate the impact of monetary transfers on economic development, the authorities must acknowledge their importance and develop a strategy aimed at stimulating their effective use.

Glossary

- **Migrants** are persons who move to a country other than that of their usual residence for a period of at least one year, so that the country of destination effectively becomes their new country of usual residence (UNPD).
- **Remittances** are the sum of workers' remittances, compensation of employees, and migrants' transfers (World Bank).
- **Workers' remittances** are current transfers by migrants who are considered residents in the destination country (IMF 2010a).

References

1. Миграция и денежные переводы. Восточная Европа и бывший Советский Союз под ред. А. Мансура и Б. Куиллина Пер. с англ. публикации Всемирного банка 2008 г.
2. The World Bank, «Migration and Development Brief 20»(World Bank Development Prospects Group, Migration and Remittances Unit, April 19, 2013), http:// siteresources.worldbank.org/INTROSPECTS/Resources/334934-1288990760745/MigrationDevelopmentBrief20.pdf.
3. Ratha Dilip. Workers' Remittances: An Important and Stable Source of External Development Finance, Chapter 7 in: Global Development Finance, World Bank. 2003.
4. De Bruyn, Tom and Umbareen Kuddus. Dynamics of Remittance Utilization in Bangladesh, IOM Migration Research Series, No. 18 2005, January available at www. iom.org; Hansen Peter. Migrant Remittances as a Development Tool: The Case of Somalila nd, Migration Policy Research, No.3, June 2004, IOM Working Papers Series; Sorensen, Ninna Nyberg. The Development Dimension of Migrant Remittances, Migration Policy Research, No. 1, June 2004, IOM Working Papers Series; Sorensen, Ninna Nyberg. Migrant Remittances as a Development Tool: the Case of Morocco, Migration Policy Research, No.2, June 2004, IOM Working Papers Series.
5. Remittance prices worldwide. The World Bank. http://remittanceprices.wordbank.org

Analysis of Automation Tools of Internal Audit Activities

Lilia Pavlova[1]
[1]Academy of Economic Studies of Moldova,
wiz_lilia@mail.ru, liliapav1@gmail.com

Abstract. This article addresses the issue of automation of internal audit activities. It is realized the analysis of existing most popular information systems for automation audit activities. The paper evaluated the difficulties that may arise in the process of automation, and also highlights factors that should be paid special attention in selecting solutions for automation of the internal audit activities.

Keywords: audit, automation, control, risk management, verification.

At the present rate of economic development there is a growing need for common and effective tools of automation the internal control and audit. Poor corporate governance can destroy the reputation and profitability of the company.

Conducting audits without the use of advanced technology is virtually impossible. According to the audit standards for the auditor may be difficult or impossible to obtain certain information for later verification or confirmation, without the use of information technology.

Auditing companies and internal audit departments often do not use specialized audit software due to its high cost and user acceptance. Specialized auditing system, such as Audit Command Language (ACL) or Interactive Data Extraction and Analysis (IDEA), are quite expensive and also require the special training to use them effectively.

Auditors in their activities quite actively used software products of general purpose, such as spreadsheets, text editors, reference and legal systems, as well as a database management system that provides for their organization of the storage and analysis information on the audited objects and processes and implementation of the conclusions.

Decision of many auditors use spreadsheets on account of their prevalence, ease of use, low cost, and user familiarity with the application is logically. It should be noted that for many auditors it is a temporary solution in the short term.

Use of information technology not only changed the internal audit tools, but also created additional challenges and opportunities for the organization of the internal audit. It should be noted that the use of software tools to automate the internal audit is aimed at re-engineering of the audit process.

For establishment and development of automation systems auditing activities contribute such preconditions as:

- High level of development of information technologies;
- Automation of financial accounting, control and monitoring functions in companies;
- A large volume of information processed and stored by the auditors;
- The pressure from stakeholders to reduce costs and increase efficiency by automation;
- The ability to process and analyze any information stored;
- Availability of demand for a more timely information on the conducted inspections and identified risks;
- Availability of clear rules and audit requirements for audit documentation forms;

- Establishment of requirements for the control of a wide range of risks (for example, the Sarbanes-Oxley Act, the legal requirements to combat fraud, and other legal and regulatory requirements).

Automation of the internal audit functions for the purpose of ordering processes or as tools for integrated audits for many companies is a necessity, not a luxury. In today's technologically complex world, exposed to constant changes, auditors' inadequate manual methods and controls, they need to be in step with the development of technology as advanced users of new tools.

Information technology as a tool for audit are not something new, it should be noted that the adoption of Sarbanes-Oxley Act promoted the integration of technologies in business and audit. The use of data analysis provides independent auditors look at processes, systems, financial and control operations to identify anomalies and trends.

Application of information technology allows preparing audit documentation, implementing a search and analysis of different relevant information, analyzing various performance and efficiency metrics. At the same time, specialized audit software provides solutions to the following tasks:

- Analysis of operations, transactions integrity, quality, validity and accessibility;
- Verification of calculations and computations;
- Comparison and processing of data from different sources;
- Planning of audits with the formation of the corresponding graphs for the staff and audit programs;
- Generalization of data and report generation;
- Storage of information about past audits, including identifying deficiencies and the dynamics of their correction.

Automation of internal audit functions should provide improved quality and quantity of information about the risks and control measures in the company, as well as more effective management. At the same time, automation systems of audit activity must meet the following basic requirements:

- The presence of certain functional structure providing covering of all processes of audit activity;
- Ensuring the relationship of data about the risks and control measures;
- Having the opportunity to separation of duties and responsibilities for the various departments and personnel with different levels of liability;
- Providing a number of access requirements to data and information security.

Functional structure of the automation systems of audit activity imposes requirements for the presence of information, technical, technological, organizational and legal maintenance. Each of these systems provides a specific methodology for the audit, for performing audit procedures and realizing of various operations. Functional of automation systems of audit activity must satisfy the needs of both audit firms and companies with a dedicated internal audit.

Functional structure of the automation systems of audit activity reflects the successive performing of complex operations on the implementation of the audit, from planning and ending with the formation of report deficiencies. It includes:

1. Audit Planning - defines an initial stage of a separate audit, as well as audits for a certain period for a specific company;
2. Detailed planning, work distribution and tracking their performance on each individual audit - study tool of audit and detailing the audit procedures. Each procedure of verification in system provides current information about responsible

reporters managing, the current status of related procedures and tasks execution results;

3. Risk assessment and management of working documents - important functionality for automation system of audit activity providing a risk-based approach to carrying out audit. Typically, risk assessment and storing the history of them is implemented in separately taken module and can be integrated with automation module of internal audit function;
4. Time management, costs and expenses related audits;
5. Preparing of audit report documentation - fixing the results of the performance of each individual procedure, identified deviations and recommendations for their correction, mapping other necessary information;
6. Reflection of operations and measures to address the identified during the audit violations and deficiencies - allow you to view the history of identifying deficiencies, as well as measures taken to eliminate or approval of decisions to adopt risks identified during the audit;
7. Storage of audit results.

Automation system of audit activity should include the possibility of obtaining the information needs of the audience and provide feedback to management and business units. The accumulated data and queries can be used to improve the effectiveness of the internal audit department.

Also significant functionality can be integration with corporate e-mail, task management, and notices (personal calendar), graphical representation of the data for further analysis and availability capabilities to manage templates and forms of documents. Possibility of integration automation systems of audit activity with already implemented and used in the company's systems and applications is also an additional advantage.

Automation system of audit activity should be integrated with the system of modeling business processes, which allows you to automatically import the data on changes in the structure of business processes. Integration with human resource management system provides continuous monitoring, control and analysis of efficiency of use of personnel time.

Today on the market there are specialized systems focused on the complex tasks associated auditing. Leading automation systems of audit activity are Audit Command Language (ACL), Interactive Data Extraction and Analysis (IDEA), Statistical Analysis System (SAS) and Statistical Package for Social Sciences (SPSS), which are widely used by both internal and external auditors throughout world. ACL, the market leader at the moment, used the big four accounting firms.

Most Russian systems for automation audit activity focused on financial accounting, including analysis of data contained in the client's data bases of accounting. Among the Russian systems for automation audit activity can be noted: AuditXP «Complex Audit», «IT Audit: Auditor», Audit Expert, «BAK», «ExpressAudit: PROF», ISKRA and many others.

In such systems, ACL data for testing and analysis are presented exclusively in read mode, which eliminates the possibility of changing the data by auditor. At the same time the risk of data changes in spreadsheets is high, because there is some space for the human factor. It should be noted that the specialized automation systems of audit activity can adapt to specific business requirements and provide support to international auditing standards, such as the Institute of Internal Auditors (IIA), Public Company Accounting Oversight Board (PCAOB), International Organization for Standardization (ISO).

ACL and IDEA system is not limited by volume of processed and analyzed data, limited only by the size of the hard disk. Both systems are also characterized by having to pass

specialized training courses for their future use, on the other hand less functionally developed systems are easier to use, allowing auditors to adapt more easily to use them in their activities.

Systems ActiveData, ProcessGene's Internal Audit solution and TopCAATs based on functionality of Excel, but contain their additional special features that facilitate the work of the auditors. The integration of these systems with Excel, as well as lower cost compared to ACL or IDEA, promotes growth in popularity of these systems.

Each system has its own set of built-in reports, report generator and dashboards that enable analysis of graphical data. Electronic working papers in automation systems of audit activity include a full database of documents that can be used to generate audit reports, annual reports, analysis of past audits.

Using the built-in reports and dashboards allows you to analyze: status of inspections and course of their execution; the course of execution of the audit plan; control over the use and allocation of staff time, resources, budget; view full history and calendar of the audit; Registry of risks and controls; monitoring the quality of performance of the working procedures.

Factors affecting the use of in automation systems of audit activity include:

- Technological factors related to the design, implementation and use of these systems - compatibility with installed and used software, easy to configure and use, prompt implementation of changes;
- Organizational factors - the availability of support from the management of the company, the possibility of staff development, reasonable cost of implementation and maintenance of the system, the availability of sufficient resources to support the system;
- Individual factors - the complexity of the control environment, the ordering and classification of data, existence of approved audit methodology, used a variety of systems and applications;
- External factors - adequate and quality support of system from a vendor, the operational implementation of the changes.

Growth of popularity of these systems contributes to focus on the fight against fraud, established in recent years. In this situation, the importances of internal audit are growing in view of focus on combating fraud programs. Motivation for the use of such systems supports the growth of qualification requirements for internal auditors, whose activities should become more effective in carrying out their tasks and responsibilities.

List of the system types of automation of audit activity continues to grow, improved. Proper use of these systems may lead to lower costs, increase reliability of auditor's actions, storage of audits stories.

System of automation of audit activities increase the rationality and efficiency of internal audit processes, allowing to customers to achieve efficiency and objectivity of audit. With these systems auditors are able to automate and control the processes of internal audit and lead more productive activities on risk management and internal control.

References

1. Cistov D.V., Dolgalev S.M., «Компьютерные программы для автоматизации аудиторской деятельности», Financial Academy;
2. David Coderre, "Efficiency through Automatition";
3. Lucashov R.V. «Формирование подразделения внутреннего аудита в российском банке», 2009;
4. Michele Matherly , Marcia Weidenmier Watson , Susan Ivancevich: "Implementing Generalized Audit Software in the Classroom", Volume 4, Number 1 2009 page 27 – 54;
5. Pankova S. V. «Международные стандарты аудита», study guide, Moscow, Economist, 2004;
6. «Sarbanes Oxley Act: key moments of creation of the internal control» // Edition PC Week/RE (527)17'2006, Author: Vyacheslav Vetrov, 16.05.2006.

Key Innovations in the Payment Infrastructures in Bulgaria

Silvia Trifonova
Department of Finance
University of National and World Economy (UNWE)

Abstract. The main objective of the paper is to analyze the TARGET2-BNB component as the key innovation in the payment infrastructure in Bulgaria made in the last few years. The current research is focused on the operational features and empirical development of the TARGET2-BNB national system component. It has been operated by the Bulgarian National Bank (BNB) since February 1, 2010. The introduction of this payment component has been followed by significant changes in the national payment system which are described in the paper.

Keywords: TARGET2 system; TARGET2-BNB national system component; Bulgaria.

Introduction

The payment infrastructures channel the flow of payments for goods, services and financial assets, and their smooth operation is a crucial prerequisite for the proper functioning of the financial system and the overall economy. In particular, given their extensive role and the large values and volumes of financial transactions they handle, any malfunctioning of these infrastructures can have negative repercussions for the implementation of monetary policy, the stability of the financial system and the currency. Currently, the payment infrastructures are exposed to a wide range of credit, liquidity, operational and legal risks.

The sound functioning of payment infrastructures is of great importance for the overall stability of the financial system in Bulgaria. The smooth operation of payment, clearing and settlement infrastructures also contributes to the implementation of the single European monetary policy.

The current paper aims to analyze the main innovations made in the payment infrastructures in Bulgaria that are critical to the smooth functioning of the domestic financial system in times of crisis. The paper comprises three main chapters: Part 1 demonstrates a brief overview of the literature concerning payment infrastructures; Part 2 is focused on the TARGET2 system of the Eurosystem and the TARGET2-BNB component in Bulgaria as the most important innovation in our payment infrastructure. The paper concludes with summarizing the results from the study.

1. Literature review on payment infrastructures

The standard setting body for payment and settlement systems is the Bank for International Settlements (BIS), Basel, and particularly the Committee on Payment and Settlement Systems (CPSS). This committee contributes to strengthening the financial market infrastructure through promoting sound and efficient payment and settlement system. In an environment of enormously increased volumes handled in wholesale (large-value) payment systems, the awareness of potential threats to systemic stability is particularly increased. The most important standards are the Core principles for systemically important payment systems (BIS, 2001a) which are deemed particularly essential for strengthening the financial architecture worldwide.

As concerns the European Union (EU) member states and their payment systems and infrastructures, first the Committee of Governors of the central banks of the member states of the European Economic Community (EEC) started to publish the so-called "Blue Book" in 1992. The Blue Book describes the functioning of the payment and settlement systems in the EEC. Since the establishment of the ECB in 1999, the Bank started to monitor and oversight the payment system instruments in all EU member states. The first Blue Book was published by the ECB in 2001. The ECB provides a comprehensive description of the payment systems and infrastructures in all EU member states. The ECB and the Euro area national central banks (NCBs) are interested in the prudent design and management of the payment systems which process the single European currency.

Many authors discuss the payment systems and particularly the large-value payment systems (LVPSs) such as Baliño et al. (1996), Bech and Soramäki (2001), Selgin (2004), Rosati and Secola (2005), Pagès and Humphrey (2005), Bolt and Humphrey (2005), Holthausen and Rochet (2006), Bech, Preisig, and Soramäki (2008), Carbó-Valverde and Liñares-Zegarra (2009), and others. Comparative studies of the LVPSs in the US and the Euro area are made by Martin (2005), Millard and Saporta (2005), Haldane et al. (2008), and others. The most important large-value payment system in the Euro area is the TARGET system (Trans-European Automated Real-time Gross settlement Express Transfer system).

TARGET system is established in 1999 as an EU-wide real-time gross settlement (RTGS) system which is used for the settlement of central bank operations, cross-border and domestic interbank transfers, and other large-value euro payments. It settles gross euro payments in real time using central bank money. TARGET2 system went live on 19 November 2007, replacing the existing TARGET system.

The ECB publishes Guidelines on a Trans-European Automated Real-time Gross settlement Express Transfer system (TARGET2) on 5 December 2012 (ECB/2012/27), amending Guideline ECB/2007/2 of 26 April 2007 in following areas: inapplicability of sanctions to non-Union banks; information sharing with regard to suspension or termination of access to monetary policy operations and the consequences of such suspension or termination. TARGET2 system is analyzed by many authors as Garber (1999), Tuncer (2007a, 2007b), De Sèze (2006), Becker (2008), Bonnier (2008), Scherer (2008), Poncelet (2008), Sinn and Wollmershäuser (2011, 2012), Fahrholz and Freytag (2012), Buiter et al (2011), Mody and Bornhorst (2012), Bindseil and König (2012), Cecioni and Ferrero (2012), Auer (2012), European Economic Advisory Group (2012), Cecchetti et al. (2012), Whelan (2012) and others.

Wide range of country reports are also published as regards the payment systems and infrastructures. For instance, Rybiński, 2007, examines the payment system development in Poland. Snellman, 2000, Jyrkönen and Paunonen, 2003, Jyrkönen, 2004, Hasan et al., 2009 and 2010, and others, investigate the payment systems in Finland. Racocha, 2004, studies the payment system's performance in the Czech Republic, while Kalckreuth et al., 2009, examine the experience in Germany. Harrison et al. 2005, Millard and Saporta, 2005, Merrouche and Schanz, 2008, and others analyze the UK payment system. De Sèze, 2006, examines the TARGET2 project and its implications for the Banque de France, while Bardinet, 2003, studies the French payment system. España and Gorjón study the TARGET2 and its introduction in Spain. The Bulgarian payment system development is discussed in the ECB Blue Book (2002, 2007b) as well as by authors like Iskrov (2010), Trifonov and Trifonova (2012), and others.

2. TARGET2 system and TARGET2-BNB national system component: operational features and empirical development

TARGET2 is based on a technical centralized platform – Single Shared Platform (SSP), developed by Banca d'Italia, Banque de France, and Deutsche Bundesbank. These three central banks operate the system on behalf of the Eurosystem. It has a decentralized structure linking together national RTGS systems and the ECB Payment Mechanism. TARGET2 participants access the SSP using the full set of SWIFT's messaging services: FIN, FIN Copy, InterAct, FileAct and Browse. In 2012 TARGET2 had 999 direct participants initiating payments on their own or their customers' behalf [Table 1]. A number of options are available to participants as regards access to TARGET2, including direct and indirect participation, addressable Bank Identifier Codes (BICs) and multi-addressee access. Direct participants hold an RTGS account on the SSP of TARGET2, with access to real-time information and control features. Indirect participants, addressable BICs and multi-addressees settle their payments via a direct participant, using the direct participant's account on the SSP. Including branches and subsidiaries, almost 57,000 banks across the world can be addressed via TARGET2. Furthermore, more than 80 financial market infrastructures, such as other payment systems, clearing houses, securities settlement systems and central counterparties, benefit from the opportunity offered by TARGET2 to settle in central bank money.

TARGET2 system is based on harmonized conditions for its participants, enhanced operational efficiency and reliability, common pricing scheme for all participants in the Euro area, wide range of tools for optimizing the participants' intraday liquidity management, as well as sophisticated business continuity and contingency arrangements. TARGET2 is an essential vehicle for the implementation of the monetary policy of the Eurosystem and has helped to create a single, integrated and stable money market within the euro area.

The following statistical data [ECB, 2013a: 8] demonstrate the leading position of the TARGET2 system in the European landscape. The system functioned smoothly in 2012, processing 92% of the total value of payments in large-value euro payment systems. The volume of payments processed in TARGET2 rose by 1.2% in 2012 compared with the previous year, reaching 90.6 million transactions, while the total value of payments increased by 3.5%, reaching a total of €634 trillion. The average daily volume was 354,185 transactions, with an average daily value of €2,477 billion [Table 1].

Table 1: TARGET2 facts in 2012

999 direct participants, 3,386 indirect participants and 13,313 correspondents	29 June 2012 – the peak in volume turnover with 536,524 transactions
Settled cash positions of 82 ancillary systems	1 March 2012 – the peak in value turnover with €3,718 billion
354,185 daily average number of processed payments representing a daily average value of €2,477 billion	TARGET2's share in total large-value payment system traffic in euro – 92% in value terms and 58% in volume terms
€7,1 million – average value of a TARGET2 transaction	100% technical availability of the SSP of TARGET2
68% of all TARGET2 payments with a value of less than €50,000 each; 11% of all payments with a value of over 1 EUR million each	99.98% of TARGET2 payments processed in less than five minutes

Source: ECB. http://www.ecb.europa.eu/

TARGET2 is legally structured as a multiplicity of payment systems composed of all the TARGET2 component systems, which are designated as 'systems' under the national laws.

On 31 December 2012 all of the Eurosystem's NCBs and their banking communities are connected to TARGET2. Other EU national central banks may join TARGET2 on a voluntary basis, making TARGET2 accessible to a large number of participants from 23 EU countries [ECB, 2013b:1]. The SSP provides the option that all users of TARGET2, irrespective of their location, have access to the same services, functionalities and interfaces at the same price.

The central banks operating system components are participating and connected to the SSP of TARGET2. The Bulgarian National Bank (BNB) as a central bank operates TARGET2-BNB, administers the National service desk of TARGET2 and is responsible for business relationships with its participants and coordination with the European Central Bank and the participating central banks.

On 27 November 2008, the Governing Council of the BNB decided to launch the project for the BNB and the national banking industry for joining TARGET2. The participation in the system allows banks to use the service "real time transfers in euro" which greatly reduces the time for processing payments in euro to banks within the European Economic Area (EEA), using the advantages of effective mechanisms for settlement and liquidity management, a modern technical infrastructure, and functionality and organization with a high level of security and efficiency. For retail payment systems, the participation in TARGET2 provides an opportunity to improve competitiveness and expand their field of activity, providing payment services consistent with the requirements of banks and their customers.

After the successful completion of all project activities to join TARGET2, the national component system TARGET2-BNB was launched on February 1, 2010. The ancillary system BISERA7-EUR, operated by BORICA-BANKSERVICE JSC, joined TARGET2 on the same date. BISERA7-EUR processes customer transfers in euro at a designated time. It is a payment system for servicing customer transfers in EUR, based on the rules, practices and standards of the Single Euro Payments Area (SEPA). It handles only SEPA payments and settles the resulting monetary obligations in TARGET2.

The BNB operates as a connected NCB to TARGET2 pursuant to a specific agreement. This agreement specifies that the connected BNB does comply with the ECB Guideline on a TARGET2 (ECB/2012/27), subject to any mutually agreed appropriate specifications and modifications. The ECB guidelines are implemented in the TARGET2-BNB National Payment Component Rules and Procedures.

According to the Law on payment services and payment systems, effective as of 1 November 2009, the TARGET2-BNB component is defined as a payment system in Bulgaria. TARGET2-BNB processes high-value and urgent EUR-denominated credit transfers. There is no minimum transaction size. In addition, TARGET2-BNB effects the final net settlement of EUR-denominated payments processed by BISERA7-EUR. The BISERA7-EUR system clears EUR-denominated payments, including SEPA credit transfers and SEPA direct debits. BISERA7-EURO's maximum value threshold is EUR 50,000. There are 31 direct participant banks in BISERA, including the BNB. The BISERA7-EUR system has 15 direct participant banks. TARGET2-BNB operates from 08:00 to 19:00 EET, Monday to Friday. BISERA operates 24 hours a day, seven days a week.

TARGET2-BNB national system component includes the BNB, 23 direct participant banks, three addressable BIC holders, and the BISERA7-EUR ancillary system for setting customer transfers in euro at a designated time [Table 2].

Table 2: List of TARGET2 participants via the national component system TARGET2–BNB

Allianz Bank Bulgaria AD	Investbank AD
ALPHA Bank S.A. – Bulgaria Branch	Municipal Bank AD
Bulgarian Development Bank AD	Piraeus Bank Bulgaria AD
Bulgarian National Bank (BNB)	ProCredit Bank (Bulgaria) AD
Bulgarian-American Credit Bank AD	Raiffeisenbank (Bulgaria) EAD
Central Cooperative Bank AD	T.C. Ziraat Bankasi - Sofia Branch
Corporate Commercial Bank AD	Texim Bank AD
Credit Agricole Bulgaria EAD	Tokuda Bank AD
D Commerce Bank AD	UniCredit Bulbank AD
DSK Bank EAD	United Bulgarian Bank AD
Eurobank Bulgaria AD	Ancillary system BISERA7-EUR operated by BORICA-BANKSERVICE JSC
First Investment Bank AD	Ancillary system BNBGSSS operated by Bulgarian National Bank
International Asset Bank AD	

Source: BNB. http://www.bnb.bg/

The following types of entities are eligible for direct participation in TARGET2-BNB:

(a) credit institutions established in the EEA, including when they act through a branch established in the EEA;

(b) credit institutions established outside the EEA, provided that they act through a branch established in the EEA;

(c) NCBs of Member States and the ECB.

The BNB may, at its discretion, also admit the following entities as direct participants: (a) treasury departments of central or regional governments of Member States active in the money markets; (b) public sector bodies of Member States authorized to hold accounts for customers; (c) investment firms established in the EEA; (d) entities managing ancillary systems and acting in that capacity; (e) credit institutions or any of the entities of the types listed above under groups (a) to (d), in both cases where these are established in a country with which the EU has entered into a monetary agreement allowing access by any of such entities to payment systems in the Union subject to the conditions set out in the monetary agreement and provided that the relevant legal regime applying in the country is equivalent to the relevant Union legislation. Electronic money institutions, within the meaning of the Law on payment services and payment systems, Article 76(1), are not entitled to participate in TARGET2-BNB.

Direct participants in TARGET2-BNB have to comply with the requirements set out in the Rules and Procedures of this national system component. They should have at least one payment module (PM) account with the TARGET2-BNB. Direct participants may designate addressable BIC holders, regardless of their place of establishment. Direct participants may also designate entities as indirect participants in some conditions.

Credit institutions established in the EEA may each enter into a contract with one direct participant that is either a credit institution or a central bank, in order to submit payment orders and/or receive payments, and to settle them via the PM account of that direct participant. TARGET2BNB has to recognize indirect participants by registering such indirect participation in the TARGET2 directory. Where a direct participant, which is a credit institution and an indirect participant belong to the same group, the direct participant may expressly authorize the indirect participant to use the direct participant's PM account directly to submit payment

orders and/or receive payments by way of group-related multi-addressee access. For the avoidance of doubt, payment orders submitted or payments received by indirect participants are deemed to have been submitted or received by the direct participant itself.

To join TARGET2-BNB, applicant participants have to:

(1) fulfill the following technical requirements: (i) install, manage, operate and monitor and ensure the security of the necessary IT infrastructure to connect to TARGET2-BNB and submit payment orders to it. In doing so, applicant participants may involve third parties, but retain sole liability. In particular, applicant participants should enter into an agreement with the network service provider to obtain the necessary connection and admissions, in accordance with the technical specifications; (ii) have passed the tests required by the BNB; and

(2) fulfill the following legal requirements: (i) provide a capacity opinion in a specified form, unless the information and representations to be provided in such capacity opinion have already been obtained by the BNB in another context; (ii) for the credit institutions, established outside the EEA, provide a country opinion in a specified form, unless the information and representations to be provided in such country opinion have already been obtained by the BNB in another context.

Applicants have to apply in writing to the BNB, as a minimum enclosing the following documents/information: (a) completed static data collection forms as provided by BNB; (b) the capacity opinion, if required by the BNB, and (c) the country opinion, if required by the BNB. The BNB may also request any additional information it deems necessary to decide on the application to participate.

The BNB should reject the application to participate if: (a) access criteria are not met; (b) one or more of the participation criteria are not met; and/or (c) in the BNB's assessment, such participation would endanger the overall stability, soundness and safety of TARGET2BNB or of any other TARGET2 component system, or would jeopardize the BNB's performance of its tasks as described into the Law on the BNB and the Statute of the European System of Central Banks and of the European Central Bank, or poses risks on the grounds of prudence. The BNB has to communicate its decision on the application to participate to the applicant within one month of the BNB's receipt of the application to participate.

The BNB is responsible for opening and operating at least one PM account for each participant. Upon request by a participant acting as a settlement bank, the BNB should open one or more sub-accounts in TARGET2-BNB to be used for dedicating liquidity. No debit balance is allowed on PM accounts. At the beginning and end of a business day, there should be a zero balance on the PM accounts. Participants should be deemed to have instructed the BNB to transfer any balance at the end of a business day to the account designated by the participant. At the beginning of the next business day such balance has to be retransferred to the participant's PM account. PM accounts and their sub-accounts are interest free, unless they are used to hold minimum reserves. In such a case, the calculation and payment of the remuneration of holdings of minimum reserves has to be governed by Council Regulation (EC) No 2531/98 of 23 November 1998 concerning the application of minimum reserves by the ECB and Regulation (EC) No 1745/2003 of the ECB of 12 September 2003 on the application of minimum reserves (ECB/2003/9). In addition to the settlement of payment orders in the Payments Module, a PM account may be used to settle payment orders to and from Home Accounts, according to the rules laid down by the BNB. Participants should use the ICM to obtain information on their liquidity position. The BNB has to provide a daily statement of accounts to any participant that has opted for such service.

The following payment orders are processed in TARGET2-BNB: (a) payment orders directly resulting from or made in connection with Eurosystem monetary policy operations;

(b) settlement of the euro leg of foreign exchange operations involving the Eurosystem; (c) settlement of euro transfers resulting from transactions in cross-border large-value netting systems; (d) settlement of euro transfers resulting from transactions in euro retail payment systems of systemic importance; and (e) any other payment orders in euro addressed to TARGET2 participants.

The types of payment orders are classified for the purposes of TARGET2. They are the following:

1) credit transfer orders;
2) direct debit instructions carried out under a direct debit authorization; and
3) liquidity transfer orders.

The payment orders submitted by participants are deemed accepted by the BNB if: (a) the payment message complies with the rules established by the network service provider; (b) the payment message complies with the formatting rules and conditions of TARGET2-BNB and passes the double-entry check; and (c) in cases where a payer or a payee has been suspended, the suspended participant's central bank's explicit consent has been obtained. The BNB has to immediately reject any payment order that does not fulfill the above requirements and to inform the participant of any rejection of a payment order. The SSP determines the timestamp for the processing of payment orders on the basis of the time when it receives and accepts the payment order.

Three priority rules exist in TARGET2-BNB such as the following: (a) normal payment order (priority class 2); (b) urgent payment order (priority class 1); (c) highly urgent payment order (priority class 0). If a payment order does not indicate the priority, it shall be treated as a normal payment order. Highly urgent payment orders may only be designated by central banks and participants, in cases of payments to and from CLS International Bank and liquidity transfers in relation to ancillary system settlement using the ancillary system interface (ASI). All payment instructions submitted by an ancillary system through the ASI to debit or credit the participants' PM accounts are also treated as highly urgent payment orders. Urgent payment orders are liquidity transfer orders initiated via the information and control module (ICM). ICM means the SSP module that allows participants to obtain on-line information and gives them the possibility to submit liquidity transfer orders, manage liquidity and initiate backup payment orders in contingency situations. In the case of urgent and normal payment orders, the payer may change the priority via the ICM with immediate effect. It cannot change the priority of a highly urgent payment order.

A participant may limit the use of available liquidity for payment orders in relation to other TARGET2 participants, except any of the CBs, by setting bilateral or multilateral limits. Such limits may only be set in relation to normal payment orders. The minimum amount of any of the limits is EUR 1 million. A bilateral or a multilateral limit with an amount of zero is treated as if no limit has been set. Limits between zero and EUR 1 million are not possible. Limits may be changed in real time with immediate effect or with effect from the next business day via the ICM. Participants may reserve liquidity for highly urgent or urgent payment orders via the ICM. By requesting to reserve a certain amount of liquidity for highly urgent payment orders, a participant instructs the BNB only to settle urgent and normal payment orders if there is available liquidity after the amount reserved for highly urgent payment orders has been deducted. By requesting to reserve a certain amount of liquidity for urgent payment orders, a participant instructs the BNB only to settle normal payment orders if there is available liquidity after the amount reserved for urgent and highly urgent payment orders has been deducted. After receipt of the reservation request the BNB should check whether the amount of liquidity on the participant's PM account is sufficient for the reservation. If this is not the case, only the

liquidity available on the PM account shall be reserved. The rest of the requested liquidity should be reserved if additional liquidity becomes available.

Instructing participants may predetermine the settlement time of the payment orders within a business day by using the Earliest Debit Time Indicator or the Latest Debit Time Indicator. When the Earliest Debit Time Indicator is used, the accepted payment order is stored and only entered into the entry disposition at the indicated time. When the Latest Debit Time Indicator is used, the accepted payment order should be returned as non-settled if it cannot be settled by the indicated debit time. 15 minutes before the indicated debit time, the instructing participant should be sent an automatic notification via the ICM. Instructing participant may also use the Latest Debit Time Indicator solely as a warning indicator. In such cases, the payment order concerned should not be returned. Unless instructing participants have indicated the settlement time in the above-mentioned manner, the accepted payment orders should be settled immediately or at the latest by the end of the business day on which they were accepted, provided that sufficient funds are available on the payer's PM account and taking into account any liquidity limits and liquidity reservations.

Settlement of payment orders can be executed by funding provided by: (a) the available liquidity on the PM account; or (b) incoming payments from other TARGET2 participants, subject to the applicable optimization procedures. For highly urgent payment orders the 'first in, first out' (FIFO) principle has to be applied. This means that highly urgent payment orders should be settled in chronological order. Urgent and normal payment orders should not be settled for as long as highly urgent payment orders are queued. For urgent payment orders the FIFO principle should also be applied. Normal payment orders cannot be settled if urgent and highly urgent payment orders are queued.

Clearing cycle details in TARGET2-BNB component are the following:

- TARGET2-BNB settles transactions individually in real-time and with immediate finality.
- 18:00 EET: cut-off time for customer payments.
- 19:00 EET: cut-off time for interbank payments.
- Transactions are processed electronically via SWIFT.
- Final settlement takes place across participant banks' correspondent accounts held at the TARGET2 SSP.

According to the BNB statistics [BNB, 2013: 37-38] during 2012 TARGET2-BNB processed 136,783 payments for €312,515 million, including 107,480 customer payments for €5,049 million: up to 9.3% in value and 34% in number on 2011. Payments by other system components to banks were 92.2% of the value and 88.9% of the number of all TARGET2-BNB payments. There were 534 average daily payments, worth an average of €1,221 million. The daily value peak was €2,322 million, with a daily numbers peak of 872. As a comparison, by 31 December 2012 the BISERA7-EUR ancillary system included 15 of 20 direct participant banks in TARGET2-BNB. During 2012 BISERA7-EUR processed 20,694 payments, worth €238 million, up to 182.3% in value and 103.1% in number on 2011 [BNB, 2013:38]. As Iskrov [2010:2] argues, with the implementation of TARGET2-BNB component and BISERA7-EUR, banks and their customers have at their disposal a reliable and secure infrastructure for processing small and large payments in euro.

Table 3: TARGET2-BNB Monthly Statistical Data for the Period 1 January - 31 December 2012

Payment type		Total payments		Daily average payments	
		Volume	Value (EUR)	Volume	Value (EUR)
1.	Payments sent by TARGET2-BNB participants	136 783	312 514 553 512	6 420	14 644 900 331
1.1.	to TARGET2-BNB participants (of which)	15 200	24 382 923 808	713	1 142 934 972
1.1.1.	customer payments	10 513	817 753 281	493	38 493 476
1.1.2.	interbank payments	4 687	23 565 170 527	220	1 104 441 496
1.2.	to participants from other TARGET2 national component systems (of which)	121 583	288 131 629 705	5 707	13 501 965 359
1.2.1.	customer payments	96 967	4 231 371 593	4 551	199 244 927
1.2.2.	interbank payments	24 616	283 900 258 112	1 155	13 302 720 432
2.	Payment instructions of the ancillary system BISERA7-EUR	4 187	1 079 646 562	196	50 751 460

Source: BNB.

Conclusion

Since its introduction the TARGET2 system has been outlined as an essential vehicle for the implementation of the single monetary policy of the Eurosystem. It has helped to create a single, integrated and stable money market within the euro area. On February 1, 2010, the BNB and the national banking industry joined TARGET2 system through the TARGET2-BNB national system component. The participation of Bulgaria in TARGET2 allows banks to use the service "real time transfers in euro" which greatly reduces the time for processing payments in euro to banks within the EEA, using the advantages of effective mechanisms for settlement and liquidity management, a modern technical infrastructure, and functionality and organization with a high level of security and efficiency. The BNB operates TARGET2-BNB component, administers the National service desk of TARGET2 and is responsible for business relationships with its participants and coordination with the ECB and the participating central banks. The connection of the national banking community to TARGET2 is an important step in Bulgaria's integration into the euro area payment system. It provides to the participating banks an opportunity to execute real-time euro-denominated transfers at an exceptionally low price. The analysis made demonstrates that the number and the value of the processed payments via TARGET2-BNB component have been increased during the period 2011-2012. Due to the introduction of this component and the subsequent changes made in the payment infrastructure in Bulgaria, our banks and their customers have at their disposal a reliable and secure infrastructure for processing small and large payments in euro.

References

1. Auer, R. (2012), "What drives TARGET2 balances? Evidence from a panel analysis", paper to
2. be presented to an Economic Policy panel; In: Cecchetti, S.G, R.N. McCauley and P.M. McGuire (2012), "Interpreting TARGET2 balances", *BIS Working Papers* No. 393, Basel: Bank for International Settlements, December.
3. Baliño, T., O. Johnson, and V. Sundararajan (1996), "Payment System Reforms and Monetary Policy", *Finance & Development*, March.
4. BIS (2001a), *Core Principles for Systemically Important Payment Systems*, Committee on Payment and Settlement Systems, CPSS Publications No. 43, Basel, Switzerland, January.
5. BIS (2003a), *A Glossary of Terms Used in Payment and Settlement Systems*, Committee on Payment and Settlement Systems. March. Basel, Switzerland.
6. BIS (2003b), *Payment Systems in the Euro area*, CPSS-Red Book, Bank for International Settlements, Basel, Switzerland.
7. BIS (2005), *New Developments in Large-Value Payment Systems*, Committee on Payment and Settlement Systems, Publication No. 67, May, Basel, Switzerland.
8. Bech, M., and K. Soramäki (2001), Liquidity, Gridlocks, and Bank Failures in Large-Value Payments Systems, Central Banking Publications, *E-Money and Payment System Review*, pp. 111-26.
9. Bech, M.L, Preisig, C., and K. Soramäki (2008), "Global Trends in Large-Value Payments", *Economic Policy Review*, Federal Reserve Bank of New York, September.
10. Becker, D. (2008), *TARGET2. Project management and change-over from a "Level 2" perspective*, TARGET2 Conference, Paris, 31 January 2008.
11. Bindseil, U., and P. König (2012) "TARGET2 and the European sovereign debt crisis", Kredit und Kapital, Vol. 45, No. 2, pp. 135–74.
12. Bolt, W., and H. Schmiedel (2009), "SEPA, Efficiency, and Payment Card Competition", *ECB Working Paper Series*, No. 1140, Frankfurt am Main: European Central Bank, December.
13. Bolt, W., and D. Humphrey (2005), "Public good issues in TARGET: natural monopoly, scale economies, network effects and cost allocation", *ECB Working Paper Series*, No. 505, Frankfurt am Main: European Central Bank, July.
14. Bonnier, V. (2008), *TARGET2. Project management and change-over from a "Level 3" perspective*, TARGET2 Conference, Paris, 31 January 2008.
15. Buiter, W., E. Rahbari, and J. Michels (2011), "TARGETing the wrong villain: Intra-Eurosystem imbalances in credit flows", *Global Economics View*, Citibank.
16. BNB, *Law on Payment Services and Payment Systems*, Adopted by the 40th National Assembly on 12 March 2009; published in the Darjaven Vestnik, issue 23 of 27 March 2009; effective as of 1 November 2009 (amended, issue 24 of 2009, amended, issue 101 of 2010; amended, issue 105 of 2011; amended, issue 103 of 2012).
17. BNB (2013), *Annual Report 2012*, Bulgarian National Bank, Sofia, May.
18. Carbó-Valverde, S., and J.M. Liñares-Zegarra (2009), "How effective are rewards programs in promoting payment card usage? Empirical evidence", *ECB Working Paper Series*, No. 1141, Frankfurt am Main: European Central Bank, December.
19. Cecchetti, S.G., R.N. McCauley and P.M. McGuire (2012), "Interpreting TARGET2 balances", *BIS Working Papers* No. 393, Basel: Bank for International Settlements, December.
20. Cecioni, M., and G. Ferrero (2012), "Determinants of TARGET2 imbalances", *Banca d'Italia Questioni de Economia e Finanza*, no. 136, September.
21. De Sèze, N. (2006), "TARGET2: from concept to reality", *Banque de France Bulletin Digest*, No. 149, May.
22. España, R., and S. Gorjón (2005), *The European Payment Landscape: TARGET2 and SEPA*, Banco de España, Payment Systems Department, June.
23. European Economic Advisory Group (2012), *The EEAG Report on the European Economy*, Chapter 2, "The European Balance-of-Payments Problem", CESifo, Munich 2012, pp. 57–81.

24. ECB (2002), *Blue Book. Payment and securities settlement systems in accession countries*, Vol. 1, Frankfurt am Main: European Central Bank, August.
25. ECB (2007a), *Blue Book. Payment and Securities Settlement Systems in the European Union: euro area countries*, Vol. 1, Frankfurt am Main: European Central Bank, August.
26. ECB (2007b), *Blue Book. Payment and Securities Settlement Systems in the European Union: Non-euro area countries*, Vol. 2, Frankfurt am Main: European Central Bank, August.
27. ECB (2007c), *Decision of the ECB of 24 July 2007 concerning the terms and conditions o fthe Trans-European Automated Real-time Gross settlement Express Transfer system (TARGET2)*, (ECB/2007/7), OJ L 237, 8.9.2007.
28. ECB (2007c), *Guideline ECB/2007/2 of 26 April 2007 on a Trans-European Automated Real-time Gross settlement Express Transfer system (TARGET2)*, OJ L 237, 8.9.2007.
29. ECB (2012), *Guideline ECB/2007/2 of 5 December 2012 on a Trans-European Automated Real-time Gross settlement Express Transfer system (TARGET2)*, (ECB/2012/27), OJ L 333, 12.12.2013.
30. ECB (2013), *Guideline ECB/2007/2 of 26 September 2013 on a Trans-European Automated Real-time Gross settlement Express Transfer system (TARGET2)*, (ECB/2013/37) OJ L 30, 30.1.2013.
31. ECB (2008), *Financial Stability Review*, Frankfurt am Main: European Central Bank, December.
32. ECB (2009a), *A Single Currency – An Integrated Market Infrastructure*, Frankfurt am Main: European Central Bank, August.
33. ECB (2009b), *Assessment of TARGET2 against the core principles*, Frankfurt am Main: European Central Bank, May.
34. ECB (2013a), *TARGET Annual Report 2012*, Frankfurt am Main: European Central Bank, May.
35. ECB (2013b), *TARGET2*, Frankfurt am Main: European Central Bank, September.
36. Iskrov, I. (2010), "Reflections of the past year", Speech by Mr Ivan Iskrov, Governor of the Bulgarian National Bank, to the Association of Banks in Bulgaria on the occasion of the Banker's Day, Sofia, 6 December 2010, *BIS Review*, No. 169, Basel: Bank for International Settlements.
37. Fahrholz, C., and A. Freytag (2012), "Will TARGET2 balances be reduced again after the crisis?", paper presented to the University of Leipzig and DIW Berlin international conference "Intra-European imbalances, global imbalances, international banking and international financial stability", Berlin, 17–18 September.
38. Garber, P. (1999), "The TARGET mechanism: will it propagate or stifle a stage III crisis?", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 51, No. 1, pp. 195–220.
39. Jyrkönen, H., and H. Paunonen (2003), "Card, Internet and Mobile payments in Finland", *Bank of Finland Research, Discussion Papers*, No. 8, March.
40. Jikang, Z., and W. Xinhui (2004), *Financial Markets Integration in the Euro area: Developments and Obstacles*, Paper presented at the 4th Meeting of the European Studies Centers in Asia: EU Enlargement and Institutional Reforms and Asia, Macau, China, April.
41. Haldane, A.G., Millard, S., and V. Saporta (2008), *The Future of Payment Systems*, first ed., Routledge International Studies in Money and Banking, Abington: Routledge.
42. Harrison, S., Lasaosa, A., and M. Tudela (2005), "Tiering in UK payment systems: credit risk implications", Bank of England, *Financial Stability Review*, Issue No. 19, December.
43. Hervo, F. (2008), "Recent developments in intraday liquidity in payments and settlement systems", *Banque de France, Financial Stability Review*, Special Issue on liquidity, No. 11, February.
44. Holthausen, C., and J.-C. Rochet (2006), "Efficient Pricing of Large-Value Interbank Payment Systems", *Journal of Money, Credit and Banking*, Vol. 38, no. 7, October, pp. 1797-818.
45. Jyrkönen, H. (2004), "Less cash on the counter – Forecasting Finnish payment preference", *Bank of Finland Research, Discussion Papers*, No. 27, December.
46. Kalckreuth, U. von, Schmidt, T., and H. Stix (2009), "Choosing and using payment instruments: evidence from German microdata", *Deutsche Bundesbank, Discussion Paper Series 1: Economic Studies*, No. 36/2009.
47. Martin, A. (2005), "Recent Evolution of Large-Value Payment Systems: Balancing Liquidity and Risk", *Federal Reserve Bank of Kansas City Economic Review*, First Quarter 2005.
48. Merrouche, O., and J. Schanz (2008), *Banks' intraday liquidity management during operational outages: theory and evidence from the UK payment system*, Paper presented at the Joint Banque de

France / European Central Bank conference on "Liquidity in interdependent transfer systems", Paris, June.
49. Millard, S., and V. Saporta (2005), "The future of payments", Bank of England, *Financial Stability Review*, Issue No. 19, December.
50. Mody, A., and F. Bornhorst (2012), "TARGET imbalances: Financing the capital-account reversal
51. in Europe", *VoxEU*, 7 March.
52. Pagès, H., and D. Humphrey (2005), "Settlement finality as a public good in large-value payment systems", *ECB Working Paper Series*, No. 506, Frankfurt am Main: European Central Bank, July.
53. Poncelet, P. (2008), *The Contribution of TARGET2 to European Integration*, TARGET2 Conference, Paris, 31 January 2008.
54. Racocha, P. (2004), *The Payment Systems and the Financial Stability*, Keynote Presentation at Conference "Payment Systems in Central and Eastern Europe", Prague, July.
55. Rosati, S., and S. Secola (2005), "Explaining cross-border large-value payment flows: evidence from TARGET and EURO1 data", *ECB Working Paper Series*, No. 443, Frankfurt am Main: European Central Bank, February.
56. Rybiński, K. (2007), *SEPA and the payment system in Poland***,** Speech at the 13th Bank Forum "Regulatory changes in European banking – outcomes for banks and the economy in Poland", Warsaw, 14 March 2007.
57. Selgin, G. (2004), "Wholesale Payments: Questioning the Market-Failure Hypothesis", *International Review of Law and Economics*, Vol. 24, no. 3, September, pp. 333-50.
58. Sinn, H.-W., and T. Wollmershäuser (2011), "TARGET loans, current account balances and capital flows: The ECB's rescue facility", *NBER Working Paper*, No. 17626, November.
59. Sinn, H.-W., and T. Wollmershäuser (2012), "TARGET loans, current account balances and capital flows: The ECB's rescue facility", *International Tax and Public Finance*, Vol. 19, No. 4, August.
60. Snellman, J. (2000), "Evolution of Retail Payment in Finland in the 1990s", *Bank of Finland Research, Discussion Papers*, No. 19, December.
61. TARGET2 (2009), *Information Guide for TARGET2 Users*, version 3.0., October.
62. Toledano, D. (2009), "Bridging the cash and securities worlds. European experience of TARGET2", *TARGET Newsletter*, Issue No. 2, Q4, TARGET2.
63. Trifonov, T., and S. Trifonova (2012), "Some Aspects of the Bulgarian Payment System Development in Time of Crisis", *IPEDR*, Vol. 54, No. 20, pp. 101-05.
64. Tuncer, O. (2007), *An Introduction to Target2*, TARGET2 INFORMATION PORTAL, 16 August 2007.\uncer, O. (2007b), *Technical Issues with Target2*, TARGET2 INFORMATION PORTAL, 24 November 2007.
65. Whelan, K. (2012), "TARGET2 and the euro crisis", presentation to the Bank of England's Centre for Central Banking Studies, 26 June.

Information and Communication Technologies - Generator of Economic Growth for Bulgarian Companies

Maya Tsoklinova
University of Forestry, 1756, Sofia, bul. "Sv. Kliment Ohridski" № 10
malenna85@yahoo.com

Abstract: This report aims to discuss and analyze the prerequisites for establishing a system of sustainable Information and Communication Technologies (ICT) and the opportunities it offers to cope the economic challenges that Bulgarian companies face in fully participation in the information society and knowledge economy. Special emphasis is placed on state policy regarding ICT. It helps building ICT as a sector, promoting the establishment of e-government and fully e-learning.

The attention is focused on the need for the management of the ICT sector to be technically literate to be able to competently manage and monitor the more important features of information and communication technologies. It is concluded that ICT is extremely important and represent key factor for achieving sustainable growth, economic development and maximization of impact of the activities of Bulgarian companies at the microeconomic level. The role of the states to create the right policy interventions, to make investments as favorable decisions and to choose the appropriate networks of partnerships to build optimal and efficient business environment for economization of its business conduct, is highlighted.

ICT are enablers for economic growth. Special emphasis is placed on their performance as a prerequisite for the expansion of all economic sectors. Adoption and proper use of ICT will lead, among other positive effects and to improve the quality of produced goods and services. The summary that the ICT industry can be seen as a sustainable resource that managed properly, achieves a significant contribution to GDP is made.

Key words: information and communication technologies, economic growth, maximization the business effect, information society, microeconomic level.

1. Introduction

The role of state in ensuring **sustainable economic growth** is to create an environment that stimulates companies to develop their **competition corporate strategy** on the basis of long-term technological advantages but not at the expense of the transitory characteristics of the market environment. Only in this way Bulgaria will provide a place among the developed countries and will be attractive center for **foreign investments**. In this way, the government represents a huge and complex organization whose activities and strategic focus can be significantly improved by the proper application of information and communication technologies (ICT). They can serve to support improvements in productivity, effectiveness management and ultimately to improve the quality of services offered to citizens. Benefits of use of ICT in government cannot be contested. But there are some concerns about its success in terms of the strategies that have to be implemented in different countries. In this report the **specific challenges** faced by developing countries in the era of high technology are identified. The key factors that are crucial for the successful implementation of information and communication technologies are presented. A descriptive framework for categorizing key factors in the **implementation of ICT in government** is reviewed. The main factors of success (incentives and enablers), and factors associated with failure (barriers and inhibitors) are

systematized. The organizational and technological benefits are also defined. Finally, the available actions to ensure economic success for all companies that invest in the application of advanced technology are proposed. With the **appearance of information and communication technologies** (ICT) and **e-government**, it becomes possible to improve the efficiency and effectiveness of internal administration and that enable government agencies to become more flexible and open to personal needs of citizens.

2. Benefits of implementing ICT for Government and Bulgarian business organizations

This report aims to present the advantages of the application of information and communications technology for the government and thus to provide a **real opportunity** for Bulgarian companies to use them as a real generator of economic growth. The introduction of e-government will allow the creation of a favorable business environment for each company. When the benefits of the introduction of **high technologies in public activities** be evaluated the investment in ICTs of Bulgarian companies will become a strategic goal for the future that would bring **financial returns and sustainable economic development** not of the country only, but also on its individual components. In this line of thought, it can be determined that the investment and implementation of information and communication technologies in favor of the governments and companies, represents a bilateral process that ensures the overall economic development and a **high degree of prosperity** .

Informatics is a bridging discipline that is fundamentally interested in the application of information technologies and systems in organizations. Then definition of informatics, government information technology can be defined as the application of information systems within the government. Therefore, this includes application of the e-government, which is made in order to **provide more efficient public services**. [1] In support of government informatics, Tapscott [2] argues that ICT leads to a "paradigm shift" entering "the era of intelligence network" and reinvention of business and government authorities. Kaufman notes that "the traditional bureaucratic paradigm characterizing by internal production **efficiency, functional rationality, decentralization, hierarchical control** is replaced with competing requirements based on knowledge, such as flexibility, organizational network, vertical or horizontal integration, innovative entrepreneurship, organizational learning. In all its totality they can lead to accelerated supply of customer-oriented services and strategy that emphasizes coordinated networking, cooperation and external customer service" all of which are supported by ICT [3] .

According to Kaul and Odedra, governments worldwide are engaged in the process of implementing a wide range of applications related to ICT [4]. Advanced in this respect, for example, are the United States, Canada, Western Europe and Japan, and less developed countries including Argentina, Brazil, India, Mexico, Kenya and Bulgaria. For all countries, the use of ICT in government leads to increased investment activity. Western countries are convinced that the **information society** will bring **economic and social benefits**. The author cites the Organization for Economic Cooperation and Development and notes that information infrastructures will stimulate economic growth, which will increase productivity, create job places and improve the quality of life. Heeks notes that there is a big difference between the application of ICT and its use in developed and developing countries. The differences consist in what extent the **various problem areas** are pressing in specific countries [5]. It could be argued that with their adequate resources and advanced technology, Western countries have an easier way to implement ICT projects in comparison with developing countries. Most developing countries, including Bulgaria, are characterized by limited computer applications

in the public sector, inadequate infrastructure and lack of skilled labor. Odedra noted that "this situation exists, not just due to lack of financial resources, but largely due to **the lack of coordination** at different levels of the effective use of technology". These uncoordinated efforts can only lead to duplication if each department performs its projects in ICT, without taking into account the compatibility within the government [6].

Based on various studies, the keys factors ICT environment can be classified, namely:

- Most projects in ICT were originally financed by donors.
- Some donations are made without prior consultation or analysis of the needs of "recipient organizations".
- Operational/running costs are incurred by the government. Funding including capital and human resources, is the final stage of the project.
- ICT budgets are insufficient, but there is a trend to increase.
- Lack of policies on ICT and master plans guide real investment
- Unstable ICT resources.

Figure 1 is a research framework that presents a way for the effective application of ICT technology. The framework defines the field of scientific research and shows how the implementation of ICT technology improves the quality of the information system. In turn, the quality of ICT systems affect tangible benefits. An ICT project is considered successful if there are real benefits realized. Quality of ICT can be evaluated after a careful assessment of the infrastructure in order to **determine the technical functionality**. For example, if the facilities are in networking of different departments would hardly become clear that successful implementation of ICT. This will include technical and user evaluation of the functional communication systems. The quality of the information system can be determined only by evaluating the information they generate. For example, if the information is for budgeting purposes, the question may be whether the information system can generate accurate and timely financial information.

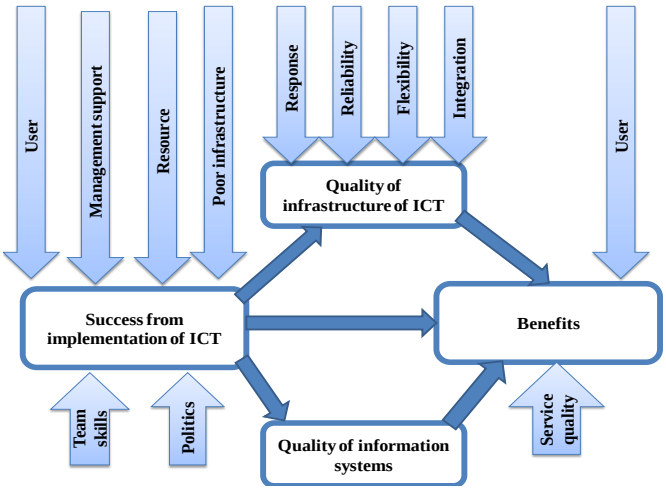


Fig. 1 Research framework

Source: Adaptation of author by Kaufman, M., Emerging Trends in Information and Communication Technologies Security, Akhgar & Arabnia, 2013, p. 662.

The benefits that are realized represent the final products, which can be used to judge the success of the entire system. The individual elements of each component are shown in the figure. A famous saw says: "If you can not plan it, do not do it". Another says: "I have never planned failure, just failed to plan". Planning is very important for the whole period of the **project implementation of ICT**. It starts with the determination of the planning of the project and strategic goals of the organization, as well as identifying specific projects in the ICT field. Within a few fixed limits, plans develop through the lifecycle project. Constraints are time and money, so every project has a clear deadline and a tight budget.

According to Moran plans fall into one of two categories: **vision without basis (budget) or budget without vision**. The problems that can be identified of vision without basis (budget) are the uncertainty of the future vision, lack of institutional vision, current position and time [7]. Identified problems of the budget without vision include questions about what problem is solved, what are the priorities and define roles and responsibilities. Maciaszek offers some models and methods for the application of ICT for planning [8]. Moreover, Aineruhanga notes that planning as a tool can help to reduce failures by setting the conditions for successful pre- application of ICT, rather than "an invasion of a comprehensive strategy for e-government, without having to first finalize the national ICT policies" [9].

Three main reasons for poor project planning for ICT in organizations can be identified. They are as follows:

- risk management has not taken place ,
- use of business systems that were not fully justified and evaluated
- lack of commitment from management .

These are the main reasons, but they should not be regarded as comprehensive and unique. This is due to a variety of environments to perform. Also, as is noted by Bannister and Reményi [1], when it comes to **complex decisions**, managers often rely on methods that do not fall within the traditional boundaries of the so-called "make rational decisions". It is observed that managers sometimes make basic decisions, guided by instinct or impulsive act of blind faith. As noted by Harindranath, although developing countries commit a large amount of economic resources for ICT, for maximizing the benefits ICT needs careful planning and coordination before implementation [10].

3. Evaluation of the opportunities for ICT in Bulgarian companies

Evaluation of investment and implementation of ICT can be determined by establishing quantitative and/or **qualitative methods**, and determine the value of ICT for the organization. The result can not be assessed as "good" or "bad" without the successful implementation of the project. Evaluation of ICT projects can be quite problematic and sometimes too subjective. It should be noted that there are common methods for the assessment of ICT that can be applied in all companies. The literature shows that the planning and management of projects in the field of ICT has a very poor performance in developing countries. However, a careful examination of the causes of failure identifies various factors whose presence or absence determines the success or failure of projects [11].

Many benefits can be achieved through the use of ICT in the management of any company aiming to reach **some economic results**. The authors Westrup sees ICT as a means to the goal of economic development, not as an end in itself. The importance of ICT is in their ability to assist the government in finding solutions to their problems. The cost of their implementation and use can be justified only if there are real benefits arising from them [12]. Among them can be referenced as follows:

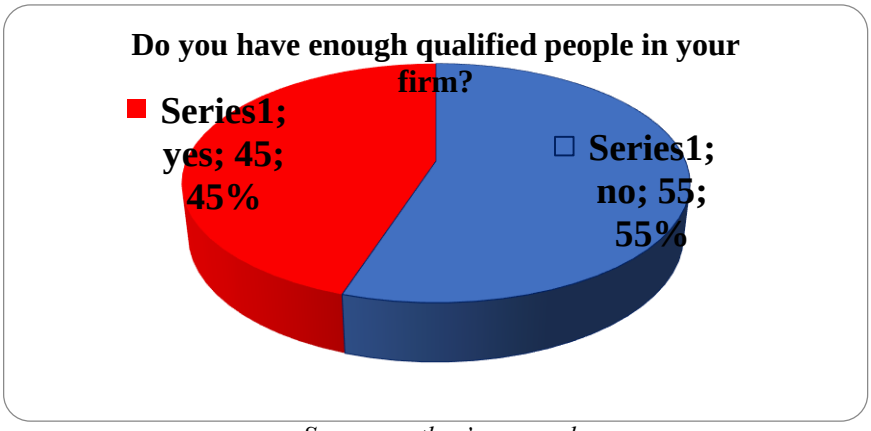
- Reduce costs.

- Quality of provided services.
- Increasing the capacity of the government.
- Improving the process of decision making.
- Transparency .
- Improved efficiency.
- Improving access to information.
- Other technological benefits - cheaper and more efficient access to storage capacities within the larger and more advanced computer systems.

4. Results and implications

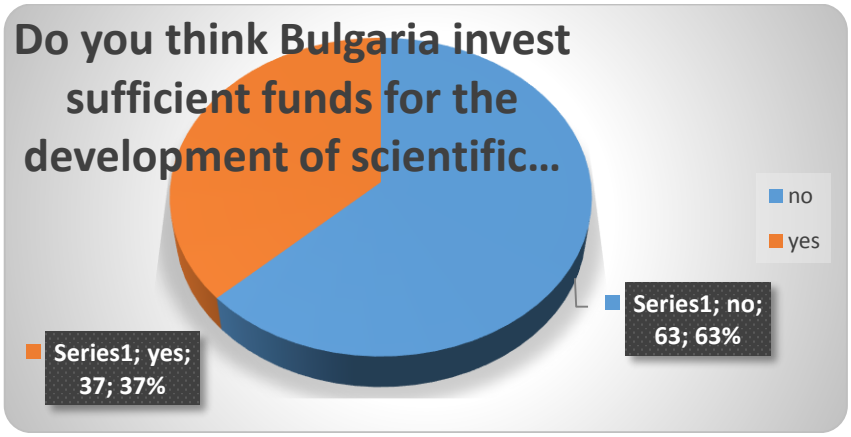
After having specified the actual benefits of the implementation of ICT in government and thus the use of high technology for the economic growth of any company, a survey was carried out among 20 managers of Bulgarian companies. The main objective of the study is to identify and highlight the benefits that ICTs provide for each business organization and identifies trends to high technology, which are a real generator of economic growth. Along with the benefits of their supply there are identified emerging issues facing the government. 15 questions were asked, grouped into 5 separate areas - trends in human resources, technological environment, development of business market environment for high technology, financial environment and fiscal environment.

To the question "Do you have enough qualified people in your firm?" 65% of managers have responded negatively. The reason for this negative result is determined by the emigration of highly skilled professionals. Extremely worrying is the fact that it is increasingly difficult for Bulgarian education system, even partially to compensate the high skilled who were depart with a new generation of experts in high technology. It is this combination of "brain drain" weak practical skills of young professionals and lack of close cooperation between universities, research institutes and businesses, that is becoming the biggest threat for accelerated technological development in Bulgaria. As one of the most important opportunities for technological development is the fact that young people in Bulgaria show exceptional interest in high technology. So, we can draw the following conclusions: Bulgarian specialists in this field have a unique combination of hardware and software skills. Unfortunately, these processes do not develop at the same rate and in other areas of high technology. **National education strategy for information and communication technologies is the desire of the authorities to improve the Bulgarian education in new technologies**, but the pace of improvement of facilities and training of teachers is not enough to be significant backlog in the past.



Source: author's research

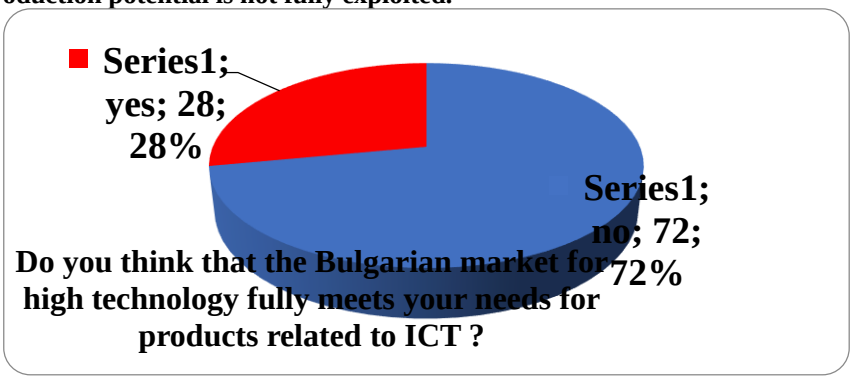
The second direction of the questions covers problem areas related to the development of the technological environment in Bulgaria, as a real opportunity to provide a favorable environment for the development of the potential of the domestic companies . As a significant disadvantage managers identify low cost of research and development, relative to GDP. To the question "Do you think Bulgaria invest sufficient funds for the development of scientific knowledge in the field of ICT?" most Bulgarian company managers have responded that public spends to stimulate research in the field of ICT are not enough.



Source: author's research

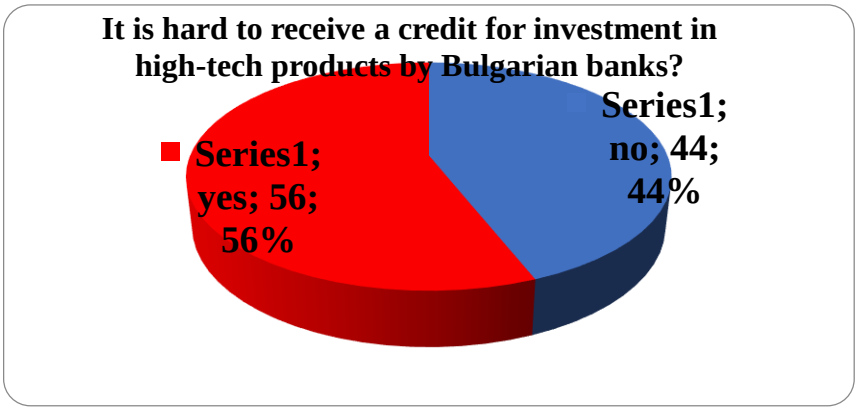
Despite the negative results, it is necessary to emphasize that the level of development of high technologies in Bulgaria can be characterized not only by the export of high-tech products, but also by the demand. In the country there is a lack technological centers and parks, in accordance with European and international standards that support business research. **A key feature of the new technologies is that they use more efficient and rational use of natural resources while producing higher quality and expensive goods.** It is reasonable to assume that in countries where the use modern technologies for processing and manufacturing industries, have achieved a good balance between energy resources consumption and the GDP .

The third group of questions covering problem areas associated with the development of the business and market environment for the development of high technology. To the question "Do you think that the Bulgarian market for high technology fully meets your needs for products related to ICT ?" 72% of respondents managers strongly determine that both the macro and micro level , the Bulgarian market for high-tech products is very limited and offers great opportunities for the development of local companies. Most of the successful high-tech companies work mainly for export, taking advantage of the opportunity to sell their goods and services to a growing global market. While Bulgarian entrepreneurs in the field of information technology, partly in microelectronics advantage of this opportunity to some degree in biotechnology existing potential is virtually untapped. In addition, most Bulgarian manufacturers of IT products with a flexible distribution and offer relatively high quality products at competitive prices. **In general it can be concluded that in all the high tech sectors production potential is not fully exploited.**



Source: author's research

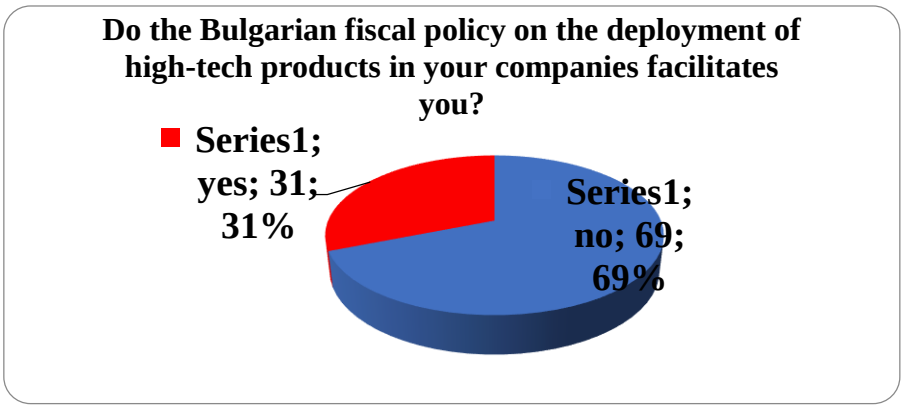
The fourth group of questions covers the advantages and disadvantages of the banking system in terms of lending, conservative and strongly restrictive lending. All these conditions have a negative impact on further technological development of Bulgarian companies. Moreover, high-tech companies have more "hunger" for funding because of the rapid turnover, more frequent replacement of equipment and the need for long-term financing. The question "It is hard to receive a credit for investment in high-tech products by Bulgarian banks ?" 56% of managers answered "Yes" , indicating that the problem of flexibility of the financial sector is not so big.



Source: author's research

For companies in the technology sector there are specific conditions for lending. In principle, this industry requires medium and long term loans at preferential terms in the initial period, but at present over 85% of all loans with maturity up to 1 year. With the exception of large companies in the technology sector for all other access to credit is very difficul. It is extremely difficult for new and emerging technology companies. Requirements of most commercial banks are for 2-3 year historical and declared profit. It is practically impossible to get a loan for business plan. Unfortunately, the access of new companies to the stock market is impossible. **Problem can be explained still by poor investment culture of company managers.** These factors are a major obstacle for innovative companies in the technology sector .

The fifth group of questions is related to the fiscal environment. Bulgarian statutory accounting rules that interfere with the tax laws are similar to principles and standards adopted by countries with developed market economy. Along with the financial stability and the reduction of corporate tax rates are positive and important trend. To the question "Do the Bulgarian fiscal policy on the deployment of high-tech products in your companies facilitates you?" 69% of managers responded that they are difficulties in this regard.



Source: author's research

The main problem areas that they claim can be divided into the following groups :

- High taxes on wages and social security contributions .
- Tax treatment of the cost of software.
- Need for further discussion of depreciation rates .
- There are no tax incentives for investment and depreciation is regarded as an expense only to the extent defined in the Law on Corporate Income Tax.
- Discrimination of exports of software products to the export of traditional products with respect to VAT.
- Unstable and unpredictable tax policy [13].

5. Conclusion

The analysis of the technological development of Bulgaria and Bulgarian companies gives reason to believe that the largely negative trends in the technology sector are broken and there is some improvement in the technological environment. Unfortunately, on the one hand, it comes after a long decline or backwardness of the Bulgarian economy in technology, on the other - can be associated with significant government initiatives in the three key elements of a viable economic environment for sustainable growth innovation, technology transfer and facilitate the creation of new business (including finance, fiscal burden of administrative barriers, etc.).

References

1. Bannister, F., Remenyi, D., Why IT Continues to Matter: Reflections on the Strategic Value of IT, Trinity College, Dublin, Electronic Journal of Information Systems Evaluation Volume 8 Issue 3, 2005, pp. 159-168.
2. Tapscott, D., The digital Economy: Promise and peril in the Age of the Networked Intelligence. London: McGraw-Hill, 1995, p. 6.
3. Kaufman, M., Emerging Trends in Information and Communication Technologies Security, Akhgar & Arabnia, 2013, p. 662.
4. Kaul, M., Odedra, M., Government initiatives in information technology application: A Commonwealth perspective. Report on the Information Technology Policy Workshop, Marlborough House, Pall Mall, London, UK: Commonwealth Secretariat, 1991, pp. 33-47.
5. Heeks, R., Building Transparency, Fighting Corruption with ICTs Available: http://www.iconnect-online.org/base/ic_show_news?sc=118&id=2363.
6. Odedra, M., IT policies in the commonwealth developing countries, Information technology policies and applications in the commonwealth countries, 1 edition. London: Commonwealth secretariat, 1993, pp. 9-35.
7. Moran, C.R., Strategic information technology planning in higher education: A new roadmap to the 21st century academy, 1 edn. Bolton, MA: Anker Publishing Company, Inc., 1998, p. 6.
8. Maciaszek, L.A., Requirements analysis and system design : developing information systems with UML. Harlow: Addison-Wesley, 2001, p.3.
9. Aineruhanga, M., Focus on the 'Kenya ICT Week', Chakula Newsletter, 2004, p. 9.
10. Harindranath, G., Information Technology Policies and Applications in the Commonwealth Developing Countries: An introduction. In: G. Harindrabath and J. Liebenau, eds, Information technology policies and applications in the commonwealth developing countries. 1 edn. London: Commonwealth secretariat, 1993, pp. 1-5.
11. Westrup, C., What's in information technology? Issues in deploying IS in organisations and developing countries. In: C. Avgerou and G. Walsham, eds, Information technology in context: Studies from the perspective of developing countries. 1 edn. Burlington, USA: Ashgate publishing company, 2002, pp. 96-110.
12. Konchev, M., Evaluation of applicability of the biometric technology in the Internet, Sofia, 2014, p. 13.
13. Беев, Ив., Fiscal effects of administrative decentralizations, Международна научна конференция на тема: „Регионалното развитие в България: тенденции и перспективи, Университет за национално и световно стопанство, София, 2012, стр., 12.

The change from ERP II to ERP III systems

Julian Vasilev[1]

[1] Varna University of economics, 9002 Varna, Bulgaria
vasilev@ue-varna.bg

Abstract. The purpose of this article is to describe several aspects of the change from ERPII to ERPIII. ERP (Enterprise Resource Planning) is a well-known concept in logistics. It is applied in software products called ERP systems. Since concepts in logistics are developing in parallel with concepts in information technology (IT), ERP systems are fast moving ahead other software systems. ERP systems are mainly intended to integrate business processes in a company. ERP II is sometimes connected with extending the functionality of ERP systems (or EPR I) towards internet integration. Other times the concept ERP II consists of modules which allow an ERP II holder to connect with some partners within a supply chain. The focus is positively towards upstream partners – mainly vendors and suppliers. The concept ERP III extends ERP I and ERP II in the direction of building new functionality of ERP systems towards integration between manufacturing enterprises and their downstream partners – mainly customers. By integrating vendors and customers in an ERP system a wide field for innovation is revealed. Manufacturing enterprises integrate information from marketplace fans with corporate data and supplier specifications.

Keywords. ERP II, ERP III, supply chain, information technology.

1. The concept ERP II and its practical implication

Companies participate in supply chains. Their participation means that they have to prove very good logistics and adoption of IT. The use of a standard ERP system allows them to integrate logistics processes – those business processes concerning material flows. ERP systems implement key logistics concepts such as MRP, MRPII and CRP. The software market of ERP systems shows a great variety form open source software solutions to proprietary software solutions. Many case studies and white papers are published. ERP systems are classified as "back office" systems.

The need for solving some information integration problems leads to extending the concept ERP I to ERP II. The need to integrate with suppliers has created a new concept "Supplier Relationship Management" (SRM) as an analogue to "Customer Relationship Management" (CRM). Integration processes with suppliers concerning new product development advances in concepts like „Product Lifecycle management" (PLM). Integration processes with suppliers focused at sending supplier information to downstream partners helped logistics managers to define the concept Advanced Planning and Optimization (APO). The concepts SRM, PLM and APO traditionally are not included in ERP I systems, but they are included in ERP II systems.

The need to reduce costs, to have fast moving goods and short time-to-market forces the logistics experts to develop new business models. Integration of information and material flows determines the adaption of a new form of collaboration. Collaborating with suppliers means sending offers from suppliers to buyers, negotiating, contracting, sending orders from buyers to suppliers, sending goods and invoices from supplier to buyers and making payments.

All these processes are existing, but few of them are covered by software products like ERP systems. ERP systems traditionally cover an Order Processing System (OPS) and an invoice issuing system. The adaption of the concept ERP II assumes that suppliers are more integrated in supply chains and they are not treated as competitors.

Since suppliers have usually a different ERP systems from the ERP system of a manufacturing enterprise new problems for information integration occur. A special vendor partnership has to be developed. The adaption of SOA (Service Oriented Architecture) is recommended [2].

Extending the base concept of ERP systems sometimes is noted as EERP (Extended Enterprise Resource Planning). EERP consists of ERP and CRM and SCM. Adapting SOA in ERP systems leads to going further and developing the concept ESOA (Enterprise Service Oriented Architecture).

2. The concept ERP III

The integration of customers (single customers as well as corporate customers) into an ERP systems lays in the essence of the ERP III concept. It may be adapted by the implementation of the last and most sophisticated achievements in IT. The focus of ERP III is not only real customers but also potential customers. Thus market information about future sales may be analyzed and integrated into ERP systems taking account of real business transactions. Extended marketing analytics are key features of ERP III. For instance gathering of information about IP addresses of internet users (See the WebApplication.IP function in Delphi XE5 [3]) show their geographical segmentation. Information about the used browser (See the WebApplication.Request.UserAgent function in Delphi XE5) may classify users by the device used to browse the web site of an ERP system or a corporate portal with integration with the ERP system.

The change from ERP II to ERP III may be done by serious software updates. These updates should be preceded by a hard work of system analyzers and system designers. Specific changes are made according to the country profile of an ERP system. Other changes are business specific. For instance companies working in the sphere of transportation logistics need extended ERP systems with modules for checking of shipments by the use of web services. The implementation of specific web services in an ERP system requires highly qualified IT staff.

Social groups and social web sites usually promote goods and services. The ERP III concept suggests the integration between ERP systems and social groups. Since the information in social web sites is highly fragmented and unstructured, it is quite difficult to integrate this type of information into an ERP system. For the time being the connection between social sites and ERP systems is not widely spread and it is not quite popular.

The ERP III concept which is the most newly known concept focuses on implementing cloud computing and GRID technologies in ERP systems. Usually ERP systems use a centralized DBMS (Data Base Management System). The use of GRID computing allows the DMBS to be separated and stored in several servers. The result is higher performance of end-user applications.

According to Ciovica and Avram [1] adapting BI (business intelligence) and KM (knowledge management) modules extend E-ERP systems.

3. Conclusion

ERP systems are widely spread and fast developing in recent years. The base concept ERP containing mainly the MRP and MRPII concepts are going further. The need for fast integration with suppliers, vendors and customers leads to creating the ERPII concept. ERPII is a concept where mainly downstream partners are integrated. The ERPII concept has the idea to integrate upstream as well as downstream partners in supply chains. The ERPIII concept is mainly influenced by the development of new IT such as service-oriented architecture, cloud computing, business intelligence and knowledge management.

References

1.Ciovica, L. and Avram C. A Grid Architecture for Manufacturing Database System. // Database Systems Journal vol. II, no. 2/2011, pp. 23-34.

2.Wood, B. SAP, ERP III, SOA — Learning Organizations through Social Media Collaboration **(**http://www.r3now.com/sap-erp-iii-soa-learning-organizations-through-social-media-collaboration).

3.Embarcadero technologies (www.embarcadero.com).

Problems of Researching the Phenomenon of "e-Commerce"

Violeta Kraeva

D. Tsenov Academy of Economics
kraeva@uni-svihtov.bg,

Petya Emilova

D. Tsenov Academy of Economics
petiaem@uni-svishtov.bg

Abstract. E-commerce is a technological, economic and social phenomenon that requires further research in three main aspects - in the terms of technology, business and society. Each of them offers potential for development.

Two main approaches in theory and practice are applied in the research and study of the electronic commerce (e-commerce): technological and behavioural, which are represented by various research fields and disciplines. From **technological** point of view, e-commerce is an object of research from: *computer science*, for which it is an application of the Internet technology; *management*, to which it provides means for improvement of the business processes based on the Internet; *systems management*, because of the influence it has on corporate strategies, value chains and structures; *macroeconomics* due to the specifics of digital electronic markets, e-money, the international e-commerce, laws, e-government, etc.

From **behavioural** point of view, e-commerce is an object of research for: *marketing*, which focuses on consumer reaction to online marketing and advertising campaigns, on brand development, customer behaviour in e-commerce sites, the ability of Internet technology to segment and target customer groups and differentiate products; electronic auctions, trust and more; *sociology* that explores the ways how Internet penetrates people's lives, using the Web as a social network and communication tool for group and other aspects; *management*, which is directed towards the study of entrepreneurial behaviour and challenges that businesses face using e-commerce, *finance and accounting*, which focus on evaluation and accounting practices of the company, whose business is e-commerce.

E-commerce is a multifaceted phenomenon, which research and study requires a multidisciplinary approach. **The purpose** of this paper is to outline the main aspects of the research and study of electronic commerce and to present important trends in its development.

Keywords. e-commerce, research approaches, trends of development

***

E-commerce (EC) is a multifaceted phenomenon, the research and study of which requires a multidisciplinary approach. The purpose of the paper is to present the main aspects of the research and study of electronic commerce and to outline the important trends in these guidelines.

1. Key aspects of the study of e-commerce

The research of the nature and taxonomy in the e-commerce have been accomplished for a long time. The contemporary vision for research and academic study of e-commerce is

the result of dynamic development. The overview of specialized major publications shows a change in the emphasis of e-commerce research. For instance:

- **Soon-yong Choi, Andrew Whinston and Dale Stahl [7]** present the following main aspects of e-commerce in "Economics of Electronic Commerce": *Technical Research* (Internet Infrastructure; Consumers' Search for Information; Electronic Payment Systems), *Economic Research* (Quality Uncertainty and Market Efficiency; Economic Aspects of Copyright Protection; Financial Intermediaries and Electronic Commerce ; Electronic Payment Systems; Business and Policy Implications), *Product Issues and Business Process* (Characteristics of Digital Products and Processes; Signalling Quality and Product Information).
- According **Urbaczweski and others** [9], studies of e-commerce can be classified into four categories: *Organizational Research* (Electronic Data Interchange (EDI)/Interorganizational System (IOS); Electronic Markets (EM); EC Diffusion and Use; EC Strategy Prescriptions.), *Economic Research* (Industry Structure, Asset Ownership, and Governance; Adoption and Distribution of IOS Value; Electronic Market Intermediaries), *Technical Research* , and others (Social; Behavioural; Legal).
- **Ngai и Wat [6]** classify the research literature regarding e-commerce in four categories: *Applications* (Interorganizational Systems; Electronic Payment Systems; Financial Services; Retailing; Online Publishing etc.), *Technological Issues* (Security; Technological Components; Network Technology and Infrastructure; Support Systems etc.), *Support and Implementation* (Public Policy; Corporate Strategy etc.).
- **Kauffman и Walden [4]** present a study of e-commerce in terms of economic analysis. They offer guidance on research for each category determined by them and define related information technology and economic theories. The authors streamline the study of e-commerce as follows:
  - *Technology Issues* (Internet Access Pricing; Intelligent Agents in EC; Search Costs and Tools; Definition of EC Standards),
  - *Product Issues* (Pricing of Digital Products; Characteristics of Internet Products; Bundling, Sharing, and Differentiation ; Adoption and Network Externalities);
  - *Business Process* (Electronic Data Interchange (EDI); Government and the Location of Transactions; Valuation of EC Technology Investments; EC Business Models);
  - *Market Issues* (Intermediation; Physical and Electronic Markets; Electronic Auctions; Network Externalities and Market Efficiency; Design of Electronic Markets; Trust);
  - *Macroeconomic Issues* (Measurement of Digital Economy; Electronic Labour Markets; Electronic Money; Taxation and Regulation of EC; Economic Transformation; Electronic Government; International EC) etc.
- **Steffano Korper and Juanita Ellis [8]** in "The E-Commrce Book (Building the E-Empire)" explore e-commerce in details predominantly in technological and business aspects.
- **David King, Jae Lee, Ting-Peng Liang, Deborrah C. Turban [1]** define the following main areas of research:
  - *E-Marketplaces*: Mechanisms, Tools, and Impacts of E-Commerce;

- *Retailing in Electronic Commerce*: Products and Services; Consumer Behaviour, Market Research, and Advertising; B2B E-Commerce; E-Supply Chains, Collaborative Commerce, and Corporate Portals;
- *Innovative EC Systems*: E-Government and E-Learning to Consumer-to-Consumer Commerce; Mobile Computing and Commerce and Pervasive Computing; The Web 2.0 Environment and Social Networks; E-Commerce Fraud and Security; Electronic Commerce Payment Systems; Fulfilling E-Commerce Orders and Other EC Support Services; E-Commerce Strategy and Global EC; Economics and Justification of Electronic Commerce; Launching a Successful Online Business and EC Projects; Regulatory, Ethical, and Compliance Issues in EC;
- *Dynamic Trading*: E-Auctions, Bartering, and Negotiations; Building E-Commerce Applications and Infrastructure

- **Efraim Turban и David King** [2] explore many new aspects of e-commerce - management aspects (performance, risk and profitability); technological aspects (Social networking, e-learning, e-government, e-strategy, Web-based supply chain systems, Collaborative commerce, Mobile commerce, Cloud computing, Crowdsourcing, RFID, f-commerce). The authors introduce new topics, such as - *Social Commerce* (Social networks, virtual worlds, social software as facilitators of social commerce, social commerce applications - social advertising and shopping, enterprise social commerce, social market research, and crowdsourcing); *Global Perspective* (global competition, partnerships, and trade).
- **Laudon K и C. Travel** [5] examined e-commerce in three main aspects: *business, technology, society*. They outline the prospects in each of these aspects and develop the following new topics: *The mobile, digital Internet platform* offer location-based services, movies, books, shopping, and payment; *Social networks* define "social E-commerce platform," and expansion of social marketing opportunities; *Online privacy*; *Internet security* .

2. Approaches to research and study of e-commerce

E-commerce is a multidisciplinary area that should be developed in collaboration with the research fields such as information systems and technologies, computer science and telecommunications, marketing, finance, supply chain management, business strategies in management, social sciences, legal sciences and others.

Two basic approaches are applied in theory and practice in the research and study in e-commerce: **technological and behavioral**. Each of these approaches is represented by various research fields and disciplines. From a ***technological perspective*** e-commerce is an object of study by: *computer science*, for which e-commerce is an application focused on Internet technologies; *management*, for which e-commerce provides means for the improvement of business processes based on the Internet; *systems management*, because of the impact that e-commerce has on company strategies, value chains and structures; *macroeconomics* due to the specifics of digital electronic markets, e-money, international e-commerce, laws, e-government, etc.

From a ***behavioral perspective*** e-commerce is an object studied from:

- *marketing*, which focuses on consumer reactions to online marketing and advertising campaigns, on brand development, consumer behavior in e-commerce websites, the

ability of Internet technology to segment and target customer groups and distinguish products; electrons auctions; trust etc.;

- *sociology* that explores ways how Internet penetrates people's lives, using the Web as social networking and group communication tool, and other aspects;
- *management*, which is directed towards the study of entrepreneurial behavior and challenges that companies face while putting into practice e-commerce;
- *finance and accounting*, which focus on evaluation and accounting practices of the e-commerce company.

3. Major trends in EC, which will define the future EC research

Some of the key trends that define the current e-commerce development [5] and hence define the directions of research in the field of e-commerce, outlined by its three main aspects, are:

Technology:

- *Mobile platforms* will continue to aggressively compete PC platforms in the future and therefore will be a research topic of present interest.
- *Cloud computing* is gaining popularity and elaborates mobile platform. They provide options for storing customer content and software on Internet servers and make them accessible from any associated client device (desktop, smartphone).
- *Big Data and diverse analytical software* for management and processing of the collected customer information is becoming more and more popular.

Business:

- Imposing new platform for e-commerce - "*social e-commerce* " based on social networks and sponsored by advertising;
- Facebook continues to grow and now has more than 1 billion active users, which are potential users in social e-commerce. Thus a new phenomenon appears: *f-commerce* that present selling with Facebook.
- Social and mobile advertising platform demonstrate tremendous growth and pose a challenge to the internet marketing.
- Demographic diversity of online users expands

Society:

- *Consumer generated content* and alliances in the form of social networks, twitters, blogs and wikis will continue to grow and form an entirely new self-published forum that involves millions of users.
- *Social networks* will continue to encourage self-revelation of consumers, while threatening privacy at the same time.
- The search for solutions continues to the problems of copyright and taxation in online commerce.

References

1. **David King, Jae Lee, Ting-Peng Liang, Deborrah C. Turban.** Electronic Commerce 2010. 6th Edition. Prentice Hall Press, 2009, ISBN:0136100368 9780136100362
2. **Efraim Turban, David King**. Electronic Commerce 2012: Managerial and Social Networks Perspectives, 7/E. 2012, Prentice Hall, ISBN-10: 0132145383
3. **Efraim Turban, Jae K. Lee, David King, Ting Peng Liang, Deborrah Turban**. Electronic Commerce 2010. 6th Edition. Prentice Hall Press, 2009, ISBN:0136100368 9780136100362.
4. **Kauffman, R. , E. Walden**, "Economics and Electronic Commerce: Survey and Directions for Research," International Journal of Electronic Commerce, Vol. 5, No. 4 (Summer, 2001), Vol. 5, No. 4, pp. 5–116.
5. **Laudon K., C. Travel**. E-commerce 2013 business, technology, society. 9th Edition, 2013, Pearson, ISBN 10: 0-273-77935-4
6. **Ngai, E., F. Wat**, "A Literature Review and Classification of Electronic Commerce Research," Information and Management,Vol. 39 (2002), pp. 415–429.
7. **Soon-yong Choi; Andrew Whinston; Dale Stahl**. Economics of Electronic Commerce. Macmillan Computer Publishing, 1997. ISBN: 1578700140.
8. **Steffano Korper and Juanita Ellis**. The E-Commerce Book (Building the E-Empire). Academic Press, 2001, ISBN: 0-12-421161-5
9. **Urbaczewski, A., L. M. Jessup, B. Wheeler**, "Electronic Commerce Research: A Taxonomy and Synthesis," Journal of Organizational Computing and Electronic Commerce, Vol. 12, No. 4 (2002), pp. 263–305.

Understanding Customer Satisfaction through a Marketing Simulation Tool in a Postgraduate IT Class

Vanya Slantcheva-Baneva

College of Management, Trade and Marketing, 1612 Sofia, Bulgaria
v.slantcheva@mtmcollege.bg

Abstract. The paper reviews the issue of understanding customer satisfaction by means of managing customer requirements during marketing strategy execution. For that purpose segment targeting decisions and marketing strategy adoptions through a Marketing Simulation Tool by Postgraduate IT students in four academic classes are represented. The central issue is inductively tackled from the perspective of a simulation-experience-sampled IT student, whose learning insights about the customer satisfaction causality are introduced.

Keywords. Marketing simulation, customer satisfaction, sales-force allocation, postgraduate IT students, qualitative assessment.

Introduction

To delineate "understanding of customer satisfaction" in the marketing strategy context is a task of multi-level perception. Marketing reveals "customer satisfaction" as an amalgam of awareness, image perceptions and loyalty. But understanding it depends on the type of both the auditory and the modus of approach. Understanding customer satisfaction by using marketing simulation tool in educational context appears to be an insightful provocation of the individual ability to recognize and appreciate causalities behind analysis of quantitative and qualitative criteria.

The purpose of this report is to review the issue of understanding customer satisfaction by means of managing customer requirements during marketing strategy execution through a simulation tool. The ultimate goal is to dispose the issue as an inductively tackled one. The induction need to be introduced from "a sampled" point of reference that is to integrate simulation-playing-behaviors of Postgraduate IT student from four academic classes.

1. Definition of Customer Satisfaction as a Qualitative Criteria for Marketing Strategy Performance

Customer satisfaction appeared to be a significant performance criterion of measuring tactical decisions. Spending on large-volume or small-volume customers, list price, channel discounts, and sales-force allocation by segment affected customer satisfaction predominantly with greatest effect on high-end switcher and lowest on low-end switcher [3]. Throughout the Marketing simulation, IT students measured the marketing strategy performance using both quantitative criteria – profitability, revenue, unit sales, and market share – and qualitative criteria referring customer satisfaction. The simulation provides IT students with a rich array of tactical decisions to retain a strong focus on marketing decision causality in terms of customer satisfaction [2].

Customer satisfaction studies explore issues related to the consumer and the firm. Empirical studies have investigated the relationship between customer satisfaction and customer loyalty, customer satisfaction and actual repurchase, which is often treated as another important means of measuring customer loyalty [4].

Although Marketing reveals "customer satisfaction" as an amalgam of awareness, image perceptions and loyalty, there are other factors that influence purchasing where satisfaction does not always play a role, i.e. lack of perceived differentiated competitors. Thus, customer satisfaction should not be the only goal for marketing practitioners [1]. Satisfaction is used as a common marketing benchmark of an organization's performance, almost to the exclusion of other issues. Customer satisfaction is the key to success and makes the emphatic statement that a satisfied customer is a repeat customer. But ideally, an organization should include other key influencers of repeat purchase in their performance reports, for example attitudinal loyalty.

2. Marketing Strategy Adoptions through a Marketing Simulation Tool

Marketing Simulation: Managing Segments and Customers® has been adopted as a learning tool for developing marketing strategies in a Marketing Management course. The simulation has been conducted in series of Postgraduate IT classes along four academic years from 2009/10 to 2012/13, and an empirical data has been collected and analyzed.

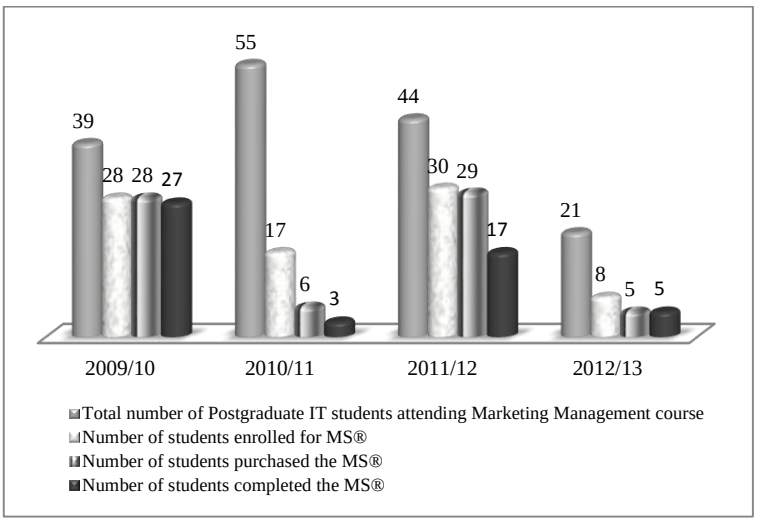


Fig. 1. Postgraduate IT student intentions to participate in online-based Marketing Simulation for four academic classes (2009/10 – 2012/13).

Throughout the simulation, Postgraduate IT students made decisions regarding marketing strategy of a manufacturer of medical motors. They determined the market segments the company should target and adopted one of three general marketing strategies of the company: "Status quo", "Price Increase" and "Commodity Play". Table 1 represents the strategic rationale of the most probable market scenarios IT students determined during the simulation. Analyzing the customer purchasing requirements, IT students made strategic decisions, on which market segments to target at and respectively which strategy to follow.

IT students predisposed themselves as conservative players because the actual point of reference in making their strategic decisions is targeting at the "Company Loyals" or the company-dominated market – Segment A. Consequently, they went in for "Status quo" strategy.

Table 2. Target segments approaches by segment definition, purchasing criteria analysis, channel conflict management, and company strategy adoption.

SEGMENT DEFINITION		PURCHASING CRITERIA	GENERAL COMPANY STRATEGY		
			"Status quo"	"Increase-Price"	"Commodity Play"
			TARGET SEGMENTS APPROACHED		
LARGE-VOLUME CUSTOMERS	SEGMENT A Company Loyals	• Product performance requirements: high "power-to-size" vs. lower "thermal resistance" ratios; • Product customization; • High level of sales-force competence and service.	×	×	
	SEGMENT B Competitor Loyals	• Product performance requirements: high level of "thermal resistance"			×
	SEGMENT C High-end switchers	• High product performance needs; • Willing to pay very high price.		×	
	SEGMENT D Low-end Switchers	• Relatively low performance levels in products; • Price sensitive.	×		×
SMALL-VOLUME CUSTOMERS		• Buy from distributors; • Highly price sensitive.	DISTRIBUTOR DISCOUNTS		
			→	↓	↑

Maintaining the "Status quo" strategy, IT students' focus was not upon making investments in product-performance attributes, but more on retaining existing customers, rather than acquiring new ones, from Segment A and Segment D through using the firm's sales-force time, deploying marketing resources for market research, and integrated marketing communications.

Any investment in product performance and an increase of the price was an attempt to pursue a position in high-end switchers' segment or an adoption of the second broad – "Increase price" – strategy (see Fig. 2). Any engagement in "Commodity play" strategy suggested a combination of (a) stopping investment in power-to-size ratios; (b) making marginal improvements in thermal-resistance attribute of the product; (c) dropping price.

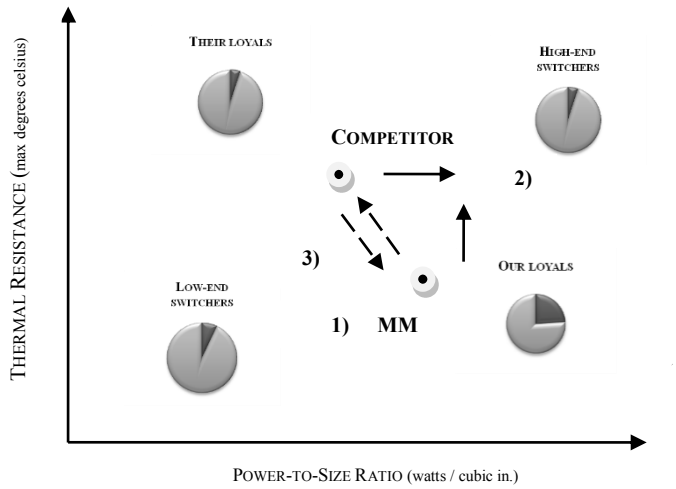


Fig. 2. Broad company's strategy alternatives against an investment in product-performance attributes. The scheme is developed by Das Narayandas [2].

All in all throughout the simulation, IT students made important decisions, regarding company product design, pricing, discount structure, marketing expenditure, sales-force size that collectively supported their strategy approaches. They played the simulation by carrying out an overall marketing strategy designed to achieve a combination of sustainable revenues and profits over 12 fiscal quarters.

3. Understanding Customer Satisfaction in Managing Customers

Considering the IT students' strategy adoptions per academic class from 2009/10 to 2012/13 empirical evidence is exposed in Table 2. The dominated strategy performed by any second IT student in learning how to execute a marketing strategy and to manage targeted segments and customers, is the "Status quo" one. Only one IT student from the 2009/10 class successfully experimented "Commodity play" strategy that was classified more like get-up-and-go attitude and ambition.

Table 2. Postgraduate IT students' strategy approaches per academic class (2009/10 – 2012/13).

GENERAL COMPANY STRATEGY	STUDENT STRATEGY APPROACHES PER CLASS							
	2009/10		2010/11		2011/12		2012/13	
	Number	Share	Number	Share	Number	Share	Number	Share
"Status quo"	17	63%	2	67%	12	71%	3	60%
"Increase-Price"	9	33%	1	33%	5	29%	2	40%
"Commodity Play"	1	4%	-	-	-	-	-	-

"Increase-price" strategy adoption was pursued by the IT students as a "high product quality – high price" strategic goal rather than setting up as a high-end market leader. Despite

of that above 30 per cent of IT students per an academic class *purposefully* invested in product-performance improvements.

A quality assessment of the results obtained in Table 2 could serve for hypothetic build-up of a sampled IT student whose simulation experience is considered as an averaged strategy approach per class. The sampled IT student was averaged out on the grounds of both the embracement of the typically conducted simulation during the four academic years, and the individual experience and yielded insights from the "Status quo" and "Increase-price" strategy execution.

Understanding customer satisfaction was taken up then upon the decision causality among customer satisfaction in, customer buying patterns, customer loyalty, and company profitability. Predominance of the "Status quo" and "Increase-price" strategy adoptions required IT students to made decisions of receiving the highest levels of customer satisfaction in Segment A and Segment D during the three fiscal years of the simulation (see Fig. 3).

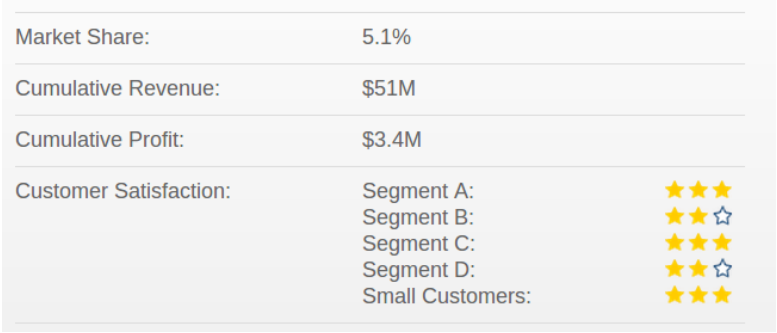


Fig. 3. Averaged results after completing the simulation by the four Postgraduate IT classes.

Assuming the recognition, the issue of "understanding customer satisfaction" is inductively concerned. The sampled IT student revealed a possible prevailing condition, which an adopted marketing strategy should take into consideration – namely, pursuing customer loyalty.

"…I realized that customer satisfaction should not be a prevailing condition when one makes operations decisions in executing an adopted marketing strategy. But from other side it could explicitly reveal existing problems with customers… I also understood that to have sales drop-offs overcome, price reductions – instead of diagnose and solve other problems with customers – rarely bring effect on it."

Conclusion

Understanding customer satisfaction through the interactive tool, *Marketing Simulation: Managing Segments and Customers®*, in Postgraduate IT classes proved to be uncomplicated and constructive task. The issue was taken up upon the appreciation of the relationships among customer satisfaction, customer buying patterns, customer loyalty, and company profitability. Assuming that recognition, the report made an attempt to dispose the issue of "understanding customer satisfaction" as an inductively tackled one. The induction was spoken from the perspective of a simulation-playing-behavior-sampled IT student. The student under consideration was averaged out on the grounds of both the embracement of the typically conducted simulation in four academic postgraduate IT classes, and the individual experience and yielded insights from the "Status quo" and "Increase-price" strategy execution.

The notion "customer satisfaction is a marketing driver" the sampled IT student perceived more or less as an assumption and a theory, rather than as an empirical proof. But undoubtedly, he stated it as an "action-reaction" determinant and measuring criterion. IT students preferably played the simulation in pragmatic manner managing segments by 1) using marketing research; 2) making investment in product performance; 3) leveraging by pricing decisions; and 4) managing channel conflict. Consequently, managing customers through using the firm's sales-force time, deploying marketing resources for market research, and integrated marketing communications IT students understand as customer satisfaction. In that respect a qualitative assessment has been made and the outcomes convincingly reveal that students call themselves "conservative players" because they *usually* apply "Status-quo" strategy, and considerable number of them enhance to "Increase-price" one. They prefer to play with a low risk so as to keep the stated market position rather than acquiring new customers. A couple of open research questions have been answered addressing the mode by which students substantiate the customer satisfaction. A prevalent student statements claim that they understand customer satisfaction as a consequence of combined decisions. Predominantly postgraduate IT students enhance product performance, manipulate the price, configure the sales force concentration and pay for market research, before all to retain the company loyals. Moreover, their playing decisions are primarily based on checking the dashboards of customer satisfaction in each quarter that notably reflects the direction – up and down – of their inputs. That feedback has made them encouraged or discouraged to continue "keep customers satisfied", which make them experience and meaningfully encounter the customer behavior.

References

1. Bennett, R. and S. Rundle-Thiele: Customer satisfaction should not be the only goal. Journal of Services Marketing, Vol. 18 (7), pp. 514-523 (2004)
2. Das Narayandas: Marketing Simulation: Managing Segments and Customers. Online Simulation, HBP for Educators, Copyright ©2010
3. Das Narayandas: Marketing Simulation: Managing Segments and Customers. Facilitator's Guide, HBS Publishing, Rev. 4048 (2010)
4. Yeung, M.C.H. et al.: Customer satisfaction and consumer expenditure in selected European countries. International Journal of Research in Marketing, Issue 30, pp. 406–416, (2013)

Mixed Economy - Quasi-Market Segments

Ivailo D. Beev,
University of National and World Economics, Sofia, Bulgaria

Key Words: Needs, interests, values, common goods, public choice, Constitution, economic system, sector models, quasi-market sector.

Introduction

The theoretical models of the mixed economic systems /these national economies differentiated with public and private sectors/ which are systematically sustained are based on the principles and the classifications, perceived by the System of the National Accounts (SNA). This way the sectors in the economy, in accordance with SNA, are defined on the base of "institutional unit" (which are two types – juridical people and households). Depending on their individual characteristics /described in three categories – "main functions in the economic process", "economic attitude" and "methods for financing"/ are spread in five institutional sectors – non-financial companies, financial companies, state management, nonprofit organizations serving the households, home economy[1].

This method of approach describes the impossibility of all of the economically important relations in the mixed economic systems to be covered and interpreted in the public and the private sectors just by analysis of the economic relations. The classification that has been established completes the tasks, connected to the measuring of the economic activity suits the needs of the national statistics, but it has weak interpretive power.

The doctrine based on the differentiation of the institutional unit determines two sectors in the mixed economies – the business sector and the sector with the sector with the public authorities. Weak and unsuccessful attempts have been made to separate the nonprofit organizations in a third sector, but the efforts in this direction have no perspective because there is no methodological or ontological based approach, which can define a sector which contains just one independent subject. At this point the level of the explanatory power of the established theoretical concepts applicable for an analysis of the mixed economic systems is inadequate. This concept also defines the goal of the current research. It premises the striving to defend one new methodological approach, which will help for creating one completely new and not controversial concept for the mixed economic systems. Accordingly the research hypothesis can be reduced to a new definition of the mixed economic systems, as trisected. Based on the precise system of criteria, the subjects and the relations can be determined, which define the three sectors in the models of the mixed economic systems. The true goal of this model is to be constructed not as an antipode of the already perceived new models in the theory of the newly introduced economic occurrences, for example transformable economic systems, globalizing markets and other occurrences and processes. Therefore the model built on the defended hypothesis, presented with the current research, appears as a more general model, which incorporates the cases of possible processes and occurrences that arise in the mixed economic models.

Developing that kind of general model of the mixed economics requires a full redefinition of the substructures and the ratified doctrines so that the prime causes for the genesis of the economic relations can be taken out, and also the reasons for the beginning of

the economic systems and their diversity. Formulate this way; the first research task appears stunningly/ quite pretentious. But actually, as Marks himself says – "even the most brilliant minds…don't notice things that are right in front of their noses. Later on when the time comes, they start to get amazed, that actually everywhere there are tracks of the same occurrences that they missed to embrace"[2].

Truly "the market" and hence "the economy" wouldn't be possible to exist if the economic subjects were not able to identify their states of dissatisfaction, by the means of which they determine their specific needs or if the economic subjects did not exert efforts for their satisfaction. I strongly believe that this statement is with axiomatic nature and this actually is the basis of the research that has been perceived. In other words, the concrete economic model is a function of the individual needs: variety of different theoretical models, that interpret one and the same model, which should be explicated by the means of the mechanisms for the satisfaction of the individual needs. The last statement defines the second research task – to interpret how the individual needs define the economic model. This predefined task requires an answer to the question – what are the processes (mechanisms), which will realize (satisfy) the needs, i.e. what is the way of satisfaction of the individual needs. If this matter transforms into a criterion for differentiation of the individual needs, then they can be divided into two big groups, which are determined as:

- Group of the "inferior" needs, which includes all individual needs, which have relatively stereotype and generally familiar model for their satisfaction
- Group of the "superior" needs, which includes all individual needs, for which a commonly known model has not been created i.e. their satisfaction happens on a individually built mechanisms.

A note has to be made, that there is a definite correlation between the two groups of needs – Depending on the level of development of the public-economic relations, the superior needs transform into inferior3, i.e. in the process of the development of the civilizations all individual needs did have superior core, but the more a definite mechanism for their satisfaction gained approval, the more they changed their nature to inferior. Simultaneously with this process, more superior needs have appeared new needs with superior nature, which show the complication of the social relations4. This transformational process can be described as one way flow as much as the civilization development can be described as such, i.e. in a certain extent an opposite connection is possible.

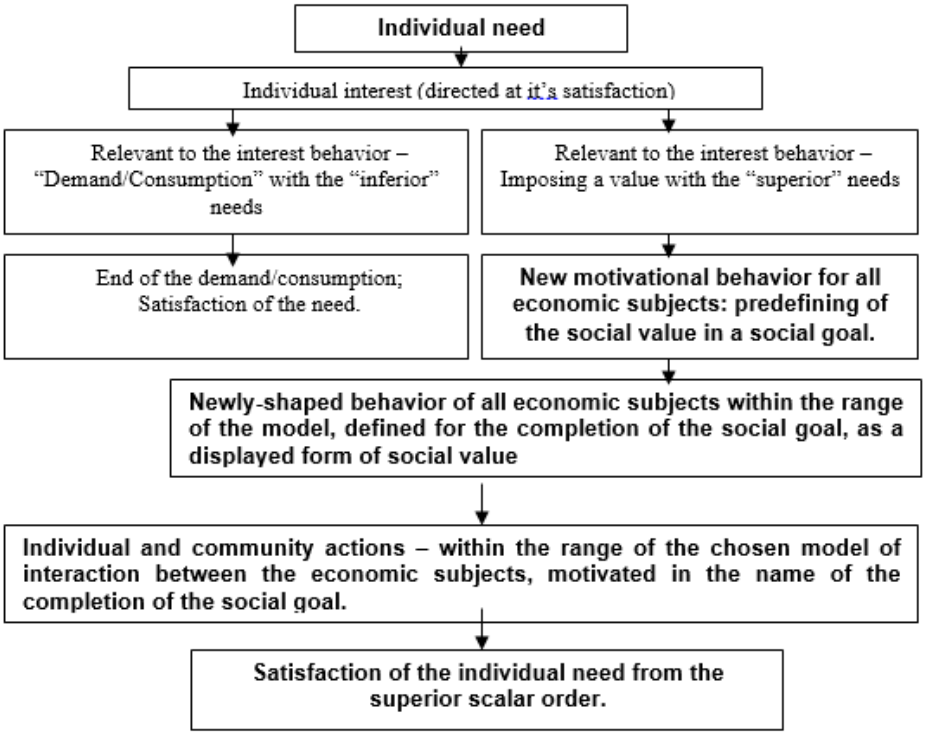
Regarding the main task – taking out the economic model via the interpretation of the needs; in the research the emphasis is put on the analysis of the superior needs, the satisfaction of which requires formulating of quality occurrences, the display of which has a normative nature. This way it leads in rules for behavior, valid for all of the economic subjects, and more importantly – it determines the parameters of the economic system i.e. it defines the nature of the relations inside the system and its' elements – the subjects and the objects for those relations. In other words – the type of the economic system and particularly the three sectors in the mixed models are determined in relation to the socially perceived values, which come from successfully imposed individual interests, directed towards the satisfaction of the individual superior needs.

Consequently the three key points in the research are analysis of the needs, analysis of the publicly (generally) determined values and the economic mechanisms, which institutionalize analysis of the economic system, defined by the values and the perceived economic mechanisms.

~ about the question for the individual needs, the Constitution and the quasi- market relations ~

If the criterion "way of satisfying the needs" is perceived as a classificatory, they can be divided into two big groups – the "inferior" and the "superior" needs. The first are defined as "inferior needs" (tbl.1 later on) are feelings that can be satisfied with direct purposeful actions on their satisfaction (i.e. there is a type of realization, which does not require previous purposeful effect on the model of the consumption). That kind of behavior has a general definition in the economic theory, which is concluded shortly with the concept "demand". The successful integration of the stereotype (standard, general) model for the satisfaction if the inferior type of needs, grants the economic subjects the freedom to shape and to motivate the behavior of other individuals or groups from the society in their attempt to satisfy their nontrivial (as a way, method, approach for the satisfaction) needs (superior).

Shaping the behavior of the other participants in the economic life, motivated by the effort toward the fullest and the most effective satisfaction of the own natural economic interests, is a need. It is determined as a "superior"5 one, but this is a single display of the subordination that has been identified and in more general case is defined (for each rational economic subject) as an individual need for increase in the personal wealth, which is a "superior" type of need.



The relevant interest, caused by this need at the Pareto – optimal economic systems, gets in a conflict with the individual interests of the rest of the economic subjects. In this case

the probability for satisfaction of the occurred need is determined by the opportunity for it to get realized successfully or to be apprehended as a common (public) choice. In economic systems that are not Pareto – optimal, the chance of successful realization of the relevant to the superior need interest is determined by the chance the behavior directed toward the satisfaction of the need, to be apprehended as permissible or if there is not any rule (criterion), such to be created. This means that in the economic order a value needs to be created, towards which the concrete behavior to be considered acceptable.

The upper analysis points out the key problem – how is the "Common good" in the Pareto- suboptimal economic systems. The thesis for cohesion of the individual interests and their simple aggregation, resulting in a common rational will6 (choice), is groundless. This understanding, interpreted more widely means that the common will should not be examined as a sum of rational individual choices. Even more – the process of the satisfaction of needs from superior type not only results directly on the behavior, defined as "demand"7, but also leads to forming values, based on which to shape and motivate the behavior of the rest of the individuals. Accordingly the common will (choice), as well as the procedures, on which it is implemented, proves to be formed under the influence of individual imposed values. Consequently in the basis of the so called "common will (choice)", there is a value, reflecting the preferences of some of the economic subjects, which have the ability to impose it (legalize it) over the rest of the members of the society.

Therefore, no matter the nature of the economic system with regard to the optimality, with the criterion of Pareto, the satisfaction of individual superior needs requires the presence of a certain value – as a minimum at least with regards to the mechanism for coordination of the wills. At the same time the economic subjects lead by their individual "egoistic" interests, by imposing the newly-formed value, they change the motivation and the shape of the behavior of all economic subjects, so that they can realize their own interest via the mechanisms of the economic relations (public and private). At this point the dynamic system of values is determined by the fundamental consensual values explicated in the norms of the Constitutions of the societies.

The current analysis applies the conclusion that the public understanding for "common good" is premised by the perceived fundamental social values – identified on a macroeconomic level, as these regarding the ownership over the resources, regarding the mechanisms of the public choice, regarding the mechanisms for the coordination of the wills etc. Those same values are defined by the text of the Constitution, which according to Giovanni Satorri8, expresses their meaning – "... she is a fundamental consensus regarding the values, which make the life of the people together in a society possible; she is a procedural consensus (i.e. who is taking the decisions, how is the choice legalized); she is also an agreement regarding the goals of the current policy (i.e. for who and how to distribute the distrainment/the premise for the wealth)"9

Therefore the Constitution recreates the fundamental value codex of the society ("consensus minimalis"10) which:

- (in a broad sense) establishes the administration of the ownership, the nature of the political system, the basic rules, based on which the relations on the distribution and the redistribution of the goods are specified (i.e. the Constitution as a form of the rules for the public choice11) and
- (in a narrow sense) Constitution – as a mechanism for the coordination of wills, as a main system creating factor, as a main "integrative technology" of the public choice12

Consequently the meaning of the choice for the combination of rules for taking decisions proves to be fundamental. From it follows the stable – because of its system nature determination (definable and predetermined) of the behavior of the economic subjects. Thus according to the principle of the voting with consentience means that everybody (according to his individual judgment) will be satisfied, i.e. each decision taken with a consentience is Pareto-optimal. But depending on the conditions of the choice (for example as it is stated transitivity in the preferences) this rule appears not to work, which forces a choice of another system of accepting decisions, and as an extreme option is "dictatorship (autocracy)". It is clear that the closer we are to the autocratic rules for making a public choice, the more is the chance for gaining a personal benefits (for the person applying the right for a public choice) to the detriment of the collective benefits. The rationally-understanding individual, or also the "rationally-purposeful subject" as defined by Marks Weber 13, puts his efforts to optimize his own upkeep when taking a decision, rendering the imposed rule for voting or the system of such rights. In this sense, this individual (economic subject) will prefer this mechanism, with which the sum of the upkeeps for finding a consensual decision and the outer upkeep (i.e. the personal losses) is minimal.

In other words – dressing in power of a certain individual or a group if individuals to take actions and make decisions on behalf of the whole majority, would have to be an object of economic analysis. To what extent this process is specified directly from the normative rules (laws, sub-delegated legislations, decisions of the Constitutional court15 etc) and to what point those rules are a supplement of the Constitutional principals, the analysis of the concrete rules for making public choice, which are regulated by the law, suggests a deep interpretive approach – it is the choice of rules, based on which, the public consensus will be made about the agreement for the way of the consumption and the distribution of the resources (including the gaining of ownership over them) and analysis over the impact of the rules (their totality) over the behavior of the economic subjects.

The democratic Constitutions tend to embrace the "…universal human values: freedom, peace, humanism, equality, justice and tolerance", which are stated in "…superior principle of the rights of the individuals, his/her dignity and security" and proclaim "… resolution to create a democratic, constitutional and social nation"16. Therefore the meaningful relations that take part in concrete constitutional texts are:

- the justice and the equality, which are defined as values.
- The nation is defined as democratic, constitutional and social

Based on this, a couple of significant conclusions follow:

- Firstly, the rules for making a public choice should be subordinate to the principle of the democracy. This means that is possible to accept decisions, which do not respond to the criterion for optimum by Pareto, with the consequences that follow – possibilities for manipulating the public choice, extracting personal benefits etc.
- Secondly, the state (the state authority) is determined as constitutional. This means that the rules for the public choice should have normative definition – i.e. the making of a public choice should rely on a legal basis. This determines the conformity with the law but it is not a premise for the expediency, or according to Weber's point – "it is not a premise for the rational-purpose. This combined with the chosen democratic way of structure (i.e. making choices a priori are Pareto-suboptimal) additionally expands, and not the opposite, the public choice to lead to violations of the principles of the optimality and the effectiveness.

As far as the democratically structured societies apprehend the specified principles and proclaim them for common values in the constitutional texts, the conclusion is applied, that the specifications of the separate economic systems does not follow directly from them. Actually the connection is indirect – with the declaring of the "constitutional state" as a constitutional principle - respectively as a common value, the different democratically structured societies regulate the economic relations in accordance with the traditions, existing in the separate national economies, by specific laws for each of the nations. Actually the combination of these laws are the needed system of criteria via which to be determined simultaneously the subjects and the objects of the relations, defining the three sectors in the models in the mixed economic systems.

On the first level of the analysis, the duplicity in the mixed economy is analyzed. It is a result from the presence of purely command (administrative) and market (competitor) combined mechanisms which form the economic activity of the economic subjects. This way the presence of the command-administrative and competitor-market economic mechanisms at the same time suggests also the presence of subjects of those relations, which can be identified as the state and the market. That kind of definition of the subjects predefines the existence of two economic lines, in the sense of subsystems of the mixed economy – the one of the "state" and the one of the "market". This way perceiving the commonly accepted terms, in the mixed economy there are two systems that can be defined by the criterion "subject of the relation" – "public (state) sector"17 and "private (non-state) sector" or defined by the criterion "nature of the economic relations" – "market sector" and "non-market sector". Therefore the presentation of the economic dichotomy as "public" and "market" sectors is not logically well sustained.

Formal-logically follows the conclusion that in order for an accurate identification of the economic subsystems as "public (state) sector and market sector" to be possible, the economic relations, in which the subjects that the "public (state) sector" and the "private (non-state) sector" are composed of, need to be realized entirely in the range of the particular sector, i.e. to be interior-sector relations. But this statement is obviously not true. There is no obstacle, and this is confirmed by the empirical reality18, for the subjects of the market relations to be publicly-legal organizations, as well as the opposite. Thinking over the last conclusion points out an important peculiarity – if a side in a market transaction is publicly-legal organization (subject), then formally the relation is defined as a "market" relation, but also has the nature of public choice that has taken place. The mechanisms of the last one are not market at all – they are subordinate to conceptually different rules – "vote", normatively (legally) defined order for making public choices and others, all ratified on a Constitutional basis. Thus in the context of the current analysis, the proclaimed in the Constitution bases and the values accepted by the public, have a concrete display, resulting in the laws. This way two main economic subjects are determined – the publicly-legal organizations and the corporative business (the private-legal organizations)19, as in this case a normative regulation of the relations between them is created, which arise because of the exchange of goods. For the analysis of these relations, a system that consists of four classification criteria is constructed – the first, regarding the nature of the delivery of the goods- then they are defined as goods that are publicly delivered and goods that are delivered privately20; the second – regarding the goods themselves – then they are sorted as "public" and "private"21; third – regarding the subject consuming the object of the delivery, then they are defined as publicly-legal and privately-legal but for conciseness and for better clarity are defined as public organizations (consumers) and private organizations (consumers) ; the fourth – regarding the nature of the normative regulation22. Also under "delivery" of a certain good, there is also included its "consumption"; i.e. the delivery a priori has a connection with the consumption of a certain

good, object of the delivery, according to the mechanism (regulation) that has been introduced. This way in the quadrant (quadrant 5) from the social-economic matrix (fig 1) which represents the relations based on the private delivery of private goods, connected to privately consumed ones means that in this case it is regarding the privately delivered and privately consumed goods, the exchange of which is characterized also regarding the normative regulation – in the current case this exchange is done freely based on the trade legalization, but is sanctioned also from the mechanisms of the public choice, the tax legalization in particular.

This way at the end all of the relations based on exchange are sorted depending on the display of the mechanisms for making the public choice.

<table>
<tr><td colspan="2" rowspan="2">Fig.1</td><td colspan="2">Characteristics of the goods</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Private goods</td><td>Public goods</td></tr>
<tr><td rowspan="4">Characteristics of the delivery</td><td>Privately delivered</td><td>Quadrant1: Transforming the public funds into incomes for private economic subjects.</td><td>Quadrant2: Realization of social gains with an element of a secondary consumption by the private economic subject – “vendor”. At this point secondary effects for the private subject remain uncalculated.</td><td>Publicly consumed</td><td rowspan="4">Characteristics of the consumption</td></tr>
<tr><td>Publicly delivered</td><td colspan="2">Quadrant3: Redistribution of the incomes/funds between pubic economic subjects.
Quadrant4: Uncalculated secondary effects for the private subject</td><td rowspan="2">Privately consumed</td></tr>
<tr><td>Privately delivered</td><td colspan="2">Quadrant5: Redistribution of the incomes between private economic subjects, with a following distrainment of wealth.
Quadrant6: Uncalculated secondary effects for the private subject</td></tr>
<tr><td>Publicly delivered</td><td>Quadrant7: Transforming the public funds into incomes for private economic subjects.</td><td>Quadrant8: Transforming of the wealth into incomes for the private economic subjects.</td></tr>
<tr><td colspan="2" rowspan="2"></td><td>Private Sector</td><td>Public Sector</td><td colspan="2" rowspan="2"></td></tr>
<tr><td colspan="2">Characteristics (subsystem, sector) of Mixed economy</td></tr>
</table>

No matter how odd it is at a first glance – the mixed nature of the economy – as consistent of corporative (private) and public sectors connected directly via the exchange (delivery) of goods with different nature. It is revealed directly when the relations are from the type “publicly-legal” organization (vendor) - “publicly-legal” organization (consumer) or from the type “privately-legal” organization (vendor) – organization (consumer), i.e. there is a paradox – the mixed nature of the economy is revealed via the simultaneous and relevantly independent realization of the economic transactions between economic subjects from homogeneous type or just “isomorphous”23. In the rest of the varieties of these relations: those from the type “privately-legal” organization (vendor) – “publicly-legal” organization (consumer) and “publicly-legal” organization (vendor) – “privately-legal” organization (consumer) the mixed nature of the economy is displayed in a specific formation – defined with the current work as “quasi-market sector”.

Reasons for this conclusion are:

1. The drastic inconsistency between “shape” and “content”24:

- “Shape” – the relations between the public and the private sectors formed at the exchange of goods, are based on the institutionalized mechanisms for making a public choice. In this sense the nature of this exchange should reflect the so called “social needs”, i.e. spending public resources, via the normatively defined

(introduced) mechanisms for that, rise the meaning of "public choice", which should be considered, that the "social needs" should be satisfied.

- "Content" – practically we see processes of redistribution, satisfying private (corporate) interests by the means of managing the process of transformation of the public resources into incomes of the "private-legal" economic subjects.

2. The system characteristic of the arising (the emerging25) of these relations, resulted from the reliability of the normative environment, as far as it is understood as an application (reflection, addition) of the Constitutional principles. About them in the last analysis was proved that they have system-creating (integrative) nature.

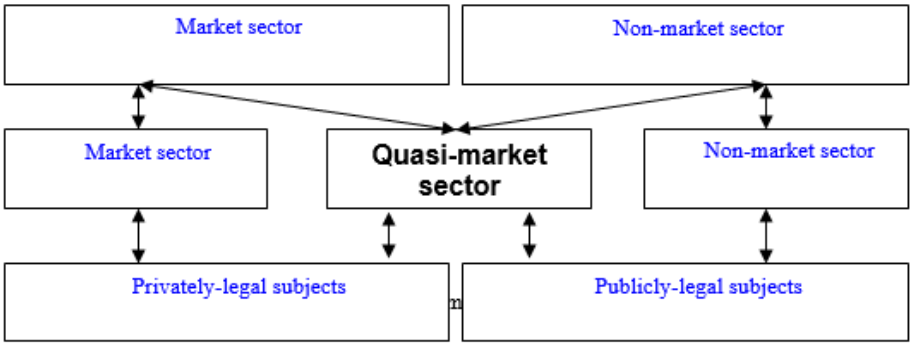
The arguments used in the defense of the hypothesis about the existence of quasi-market sector in the mixed economy, which predefines it as three sectors, is defined as "systematic proof"27. Concluding the arguments at this direction, they are reduced to defining the term "system"28 (respectively subsystem), defining the "system-building function"29, identification of the elements ("subjects" of the subsystem) and the definition of the relations30 between them. The nature of those relations- defined as "quasi-market", gives also the definition (the name) of the third sector in the mixed economy – "quasi-market sector"

In the analysis of these relations (restrictive condition for the analysis) is assumed that the individually consumed privately delivered good by the publicly-legal subject creates public consumption (i.e. public consumption is involved). The reasons for this statement are revealed not only connected to SNA32 (i.e. the reasons with methodological nature) as for basis is perceived the methodological mark that the expenses for this private delivery of the individually consumed private good will be accounted in the macro – statistics as collective consumption, since they are financed by the national budget. Alternatively as a result from the formal analysis – and exactly that the publicly-legal organizations are constructed on the base of procedure order, signed in regards with the Constitution. Therefore they express the public interests just by their origin. Then the expenses that they make for satisfying of the individual consumption via the private delivery of private goods assumingly should be taken as a priori approved by the society and following its interests. In other words, public consumption has been made, expressed in a concrete form in the individual actions for the publicly-legal economic subject, made via the direct choice of an administrator (employee etc).

The subjects of this economic interaction are the same privately-legal and publicly-legal organizations, which are subjects accordingly in the private and the public sector, but they are the same diverse economic subjects, which form inter-sector relations, creating this way a completely new type of interaction. The stable systematic characteristics of it, gives a sectored characteristics of the quasi-market segments.

In this sense, the mixed economy should be interpreted as three sector one i.e. defined by the criterion "nature of the economic relations", the sectors in the mixed economy are defined as "market sector/quasi-market sector/non-market sector".

This model has a definite specific in regards to the definition of the economic subjects functioning in the separate sectors, which applies special dynamical nature, by the means of which the model is defined as "open"34. As a type, the economic subjects, participating in one of the three sectors are constant – they are either publicly-legal or privately-legal organizations. As concrete economic subjects – there are no obstacles for them to change constantly or to have at the same time dual characteristics. This is possible because of the selected definitive criterion – "nature of the economic relations". This conclusion is visualized on figure 3.



The dynamic nature, the "free"-ness of the model requires a special analysis. If we

presume that the economic subjects are "restricted: in the specific sector, i.e. the privately-legal to function just in the private sector and the publicly-legal - only in the public sector. Then hypothetically (just conditionally) there will be a couple of effects at hand:

- Firstly, the quasi-market sector should disappear, since there are no relations between them, which can determine it, despite the presence of potential subjects.
- Secondly, with a certain delay (lag), and particularly depending on the period necessary for the depletion of the collected (in previous periods) budget resources (stock, reserve) – the public sector should also disappear because the lack of budget (financial) funds makes the economic transactions between them impossible.

As far as the existence of the public sector in the mixed economics is reliable to the taking away of financial resources from the private sector, and as far as the same resources, on a following stage are transformed into income of the privately-legal economic subjects by the public (collective) consumption, to that extent the hypothesis, about the opportunity in one mixed economic system the different economic subjects to function separately just in the specified sector, is completely groundless.

In other words the two-sector model of the mixed economics, a priori suggests the existence of relations between the publicly-legal and the privately-legal economic subjects – i.e. it proves that the assumption itself about the existence of a public sector imposes the assumption (requires the presence) of relations between the subjects in the separate sectors. These relations, however, emergent by their nature (i.e. ones that change the role, the function and the properties of the system, so that the system role, function and properties are not identical to those of the elements-relations that build it) cannot a priori be determined as quasi-market relations. In order to be transformed into such, the presence of mechanisms for public choice with quasi- market nature36 is required. This is exactly like the mechanism of the law for the social orders.

This occurrence is new – it is no more than 10 years old37, despite that theoretically separate "feelings" about the presence of this potential38 can be discovered in the works of the representatives from the 3rd and the 4th decade of 20th century. The analysis of the relations, presented figuratively in the "social-economic matrix" proves that all relations transforming public resource into private income have the necessary potential to transform into quasi-market

relations. In this current work, those relations are presented as the relations of the market destruction. Then reasonably appears the question about the moment, when this potential is realized – i.e. the moment, after which there will be not only normatively regulated relations between the publicly-legal and the privately-legal subjects, but quasi-market relations between them. In accordance with the upper analysis, this moment can be determined in the moment when the rules for transformation of the public resources into private incomes (i.e. the rules of the public choice) gain the stability quality. In the modern times it is a result from the stability of the normative environment, which is built in a particular manner by the Constitutional principles – principles that impose the concept for the public goods, but which in the end are nothing more than successfully imposed individual values.

This theoretical conclusion (hypothesis) is lead out by the author's observations over the empiric economic reality. With the promulgation of the Constitution of the Bulgarian Republic since 1991, the nature of the economic system is determined as "mixed", done by the criterion "form of ownership". Therefore by allowing the private initiative (in this case should be understood as initiative of the privately-legal economic subjects) the system characteristic "relations" between the various economic subjects is brought in. Those relations are determined as "potential", or also relations of "destruction", which six years later, in 1997, become the base for the transformation of the two-sector mixed economy into a tri-sector one, by the nominal regulation by the law for the State and Municipal orders (from the same year). This law exactly leads in the pseudo-market mechanisms for spending the public resources – i.e. becomes the basis for the relations, forming the quasi-market sector.

And if this analysis gives any reason for talking about a tri-sector model of the mixed economics here, in Bulgaria, then there should be added a very important specification, that this occurrence does not have a local nature. As far as it was proven that the arising of the quasi-market relations is determined by leading in of the pseudo-market mechanisms for public choice, which are synthesized (unified) in a common legal text – the law for the public orders (LPO), to that extend justified to trace the presence of such legal regulation for spending the public resources in other economic systems. Additionally the passage of LPO imposes the necessity of unifying (harmonizing) the normative regulation regarding the directives of the European Community "regarding the state (public) orders and the Agreement on the state orders between the participating countries from the World Trade Organization"41.

In a comparative plan there are three approaches when regulating the public orders42:

- "Privately-legal approach" – taken a line in Germany, Denmark, Great Britain, Netherlands and Ireland. Considers the public orders as a type of deal, with which the norms of the civil law are applicable. Therefore the approach that has been brought in means that the contracting authority for the public orders (i.e. the publicly-legal organizations) is given sovereign (independent) status when determining the executive privately-legal organization – same as with the privately-legal subjects, which depending on their own(independent, detached) interests and criteria to have economic relations with any economic subject.
- "Purely publicly-legal system" – the ground is regulated by the norms of the administrative law. As far as this as a legal sector, serves the arising specific relations between the legal subjects in the process of organizing and realizing the executive actions by the state (public) authorities and to that extent it is defined as a complete opposite of the "privately-legal approach". It has taken a line in France and Spain.
- "mixed system" – this third approach combines elements from the "privately-legal" and the "publicly-legal" regulations of the approaches (the

methods, mechanism) for spending the public resources (i.e. implicitly – making a public choice). This approach exactly has been adopted in the Bulgarian economic system, as well as in Belgium, Italy and Portugal.

In the Bulgarian economic system the "mixed system" is normatively established for regulating the public orders since assigning a public task (implicitly – making a public choice and transforming the public resources in private incomes) is always as a result of the display of the factual composition44 (consecutively displaying occurrences, bound by the empiric reality) including two groups of elements45: respecting a procedure (i.e. displaying the administrative nature with the choice) when determining the executor and signing a contract by the rules of the Trade law and the Law for the obligations and contracts (i.e. displaying the market nature with the choice). Therefore in the Bulgarian economic system (which is with a "mixed" nature – because of the various subjects, categorized by the criterion "forms of ownership") there is a mechanism (more concrete – the LPO) for making a public choice, which transforms the potential ("destructive") relations in the quasi-market systems.

~ Conclusion ~

The satisfaction of the specific individual needs, defined as "superior" forms economic interests, the realization of which imposes a social perception of values, as via them the realization of the economic interest becomes possible. This way the newest occurrences in the economic systems can be explained – the regulation of the "transition", the globalization of the markets, military conflicts etc.

But the more considerable contribution of this approach in the analysis of the etymology of the economic systems is the introduction of the concept for the combination of criteria that make the defining of the mixed economies to a third sector possible, which is defined as quasi-market sector. The quantity of the financial resources of the public sector, transformed into incomes for the privately-legal economic subjects, discovers the value of such formulations when interpreting the economic dynamics in the mixed economies.

The examined occurrence is new, supported by the fact that there is no specialized literature that can interpret it. I do hope that with the identification of this occurrence in this current work, a new beginning for series of new researches in this area will start.

~ notes ~

1. More on the topic about SNA – check the collective of authors „Курс социально-экономической статистики", publisher. „Омега – Л", Moscow 2007.
2. Marks in his letter to Fr. Engels dated 25 March 1868 (My translation)
3. The hypothesis is that the gaining of experience and knowledge about the etymology of the needs and the ways for their satisfaction minimizes the individually needed efforts for overcoming the realized feelings of dissatisfaction. This results in the personal development, also in the social development by ratifying a certain vital minimum, and later on – in social standard. The achievement of that in a quick way, suggests ascension to a new quality level, by which the cyclic recurrence in the development of the personal interests and the public values is formed.
4. The process of arising needs from a higher level and their satisfaction by priorities is defined by Maslou, as organizing the needs in "hierarchy of the relevant power" But in his analysis the missing thing is the interpretation of the values as a power motivating the appearance of needs from a higher level. Check Maslou, "Motivation and personality", publisher "Kibea", Sofia, 2001 pages 80 – 94
5. The definitions "superior" and "inferior" as categorization terms to the needs was brought in, and is terminologically allowable:

- most oftenly used in the economic literature is the categorization of the needs by Maslou. He differs two main types of needs – "instinct-oid" (also known as basic) and "meta-needs" which are devided into five groups – "physiological, for security, belonging, valuation and self-updating. Check Maslou "Motivation and personality", publisher: "Kibea", Sofia – 2001; pages 80 - 143.
- in the economic theory the term "inferior" is referring to the goods. But in this work the term is applied to the needs. In this sense there is no formal logical contradiction. It is inadmissible the mixture of using the terms as synonyms, despite that there is no barrier for the "inferior" goods to satisfy "inferior" needs. This of cource does not mean that the "inferior" goods can be indentical to the "inferior" goods – it is clear that the goods and the needs are from a different quality level and also there is no barrier by consuming "luxurious" or "normal" goods to satisfy "inferior" needs.

6. This actually is the basic hypothesis, based on which the term "social wealth" is determined. Check Arrow Kennet J. "Social choice and individual values", Yale University press 1963, page.17.
7. The needs coming from the economic basis of the public community form a behavior , indentified in the economic theory as "Demand", no matter the nature of the goods – commodities of Gifen, low-value commodities, normal and luxeous commodities or the type of the needs – basic, social or "ego" (the way that Eibraham Maslou categorized them)
8. See Giovanni Satori, "theory of the democracy" – publisher SU "sv Kliment Ohridski", Sofia, 1992, pages 102 – 105
9. There again; the text in italics – mine. Check also Manov, "strategy for the development and the structural reorganization in the Bulgarian economy until 2025" publisher: UNWE Sofia 2004 – page 13 "…The economic policy has always been an interest of the governing political power". Therefore the current economic policy presented under the idea "the Common good", at the end runs towards it, meanwhile realizing privately-corporative interests of the politic parties.
10. The Term "consensus minimalis" is defined as an agreement regarding the main values, presented as "Common good", which give reason to the people to live together in one society so that they are in a subordinate relation compared to the authority of the political power and at the same time to be able to complete their own interests – alone or with the coordinated actions from the rest of the participiants of this society.
11. See page 107, "The public sector economics" Brown&Jakson, publisher PSSA, Sofia 1998г.
12. The vision for the Constitution as a "common integrative technology for the public sector" belongs to the author (my own). Therefore it does not bind the opinion of third parties with it.
13. Check Maks Weber, "Sense and value", publisher: "Critique and humanism", Sofia 1998, page 168
14. "The power" according to Maks Weber means that the opportunity to be realized inside the system of collective or social relations, the own will – despite an eventual resistance (disagreement) and to the disadvantage of the rest. Check Maks Weber, "Sence and value", publisher: "Critique and humanism", Sofia 1998г., page 190-198; check also Alfred Shuller and Hans Gunter Kryusselaberg "Анализ экономических систем", изд."Экономика", Moscow 2006 page 215. б.а.
15. The different legal doctrines settle the questions regarding the Constitutional control in a different way. In the Bulgarian law doctrine, this problem is solved with the creation of a Constitutional court. The consequent analysis is based on the facts from the Bulgarian legal realit.
16. The Constitution of the Bulgarian Republic – promulgated SN ed.56 from 13 July 1991г., and added, ed.85 from 26 September 2003г. и ed.18 from 25 February 2005.
17. The author is reports the presence of a difference in the range of the terms "public sector" and "State sector", but as far as the current work, the rules for the public choice are examined, and until what extent this mechanism is common for both the sectors, so using the words as synonyms is justified and permissible.
18. Check the presented in the introduction model of social-economic interaction and it is "the social-economic matrix"
19. Check the upper note.
20. We are abstracting from the median case – i.e. when the deliveries are with mixed characteristics, because in a certain extent they can always be equalized to one of the both cases
21. For them the same presumption is available (please check the previous note)

22. Here are taken in mind the relations for exchange between the economic subjects that realize them, taking advantage from the "gaps" in the normative regulation.
23. "isomorphia" in a literal translation from Greek means "equivalent", check authors "Bulgarian interpret dictionary" publisher "Science and art", Sofia 1994. Inserting this term in the analysis, the author marks a group of economic subjects, arose on a general principle. In this sense two main groups subjects are divided – subjects of the civil law and those to the public law, which directly correlates with defining the economic subjects, as such with privately-legal nature and those with publicly-legal nature.
24. Check Mamedov "Смешанная экономика: двухсекторная модель", publisher "КноРус", Ростов-на-Дон 2001г., pages46-48.
25. The emerging means that "the characteristics, the features and the role of the system (subsystem) are not identical to the characteristics, features and the role of the elements that are building it. The definition by Manov, Vassil "Prognosis and Planing the development and the functioning if the economic systems" – publisher UNWE – "Economics" department, Sofia, 2001, page 31, the text in bold is mine; check also Mirkovich, Kamen "Utility and value", publisher "Trakia – M".
26. Check also Velento and Eliseev "Теория экономического право", publisher."Книжный дом", Minsk 2006.
27. Obviously that ("the systematic proof") is with purely theoretical nature which imposes its empiric argumentation, but before that a second theoretical proof of the hypothesis about the tri-sectoring in the mixed economics – the so called "semantic proof".
28. The term "system" is defined by Bertalamfi as a "complex from elements, that interact with each other" or according to Holl and Feidgin as a "combination of objects and connections between them". Russel Akoff gives a more detailed definition about the term "System" by defining it as a "combination of two or more elements, that cover the following three conditions:
    - the behavior of each element influences the behavior of the whole.
    - the behavior of the elements and their impact on the whole are interdependent.
    - If three groups of elements exist, each of them influences the behavior of the whole and none of them do not have that impact by itself. In other words, the elements of the system are connected in between in a way that from them no independent groups can be created.

    More thoroughly about the definition about system – check Manov, Vasil "Prognosis and Planing the development and the functioning if the economic systems" – publisher UNWE – "Economics" department, Sofia, 2001 pages 26-33
29. The elements of the economic system – reviewed as "subjects" or like "relations", they arise and function with correspondence to the Constitution. Therefore in is correct to define the Constitution as a "attribute for creating systems".
30. understood as relations of determination.
31. "publicly-legal" economic subject means each social organization, spending resources no matter if it has incomes from it's activity. This definition to be given (to be brought in) later in the text!!
32. It is the System for National Accounts.
33. For the concrete analysis doesn't matter if the financing is 100% state or mixed i.e. one that should be made by publicly-legal organization, which has also an incomes part in it's budget (to form incomes from it's activity)
34. The models from the "free type" allow a certain variability, even indefiniteness regarding all the changes and determinants. In the models from this type the analysis is concentrated onto the main relations that also allow being preformed. Fully and in details, this type of mathematical shaping in the harmonious part of the current research.
35. Check check Manov, Vasil "Prognosis and Planning the development and the functioning if the economic systems" – publisher UNWE – "Economics" department, Sofia, 2001 page 31
36. This exactly adds the characteristic of determination at the relations, so that it becomes possible to speak about separating of a different subsystem, characterizing itself by them.
37. In the Bulgarian legal system, and with that in the economic relations, a similar mechanism for making public choice in the conditions of mixed economics was lead in for the first time with the law for the State and Municipal orders from 1997, which was canceled by the LSO since 1999 which was canceled again by the LSO from 2004. The last one is the actual legal instrument at the moment.

38. To be understood as presence of normatively-regulated relations between the economic subjects from aspectual type i.e. Privately-legal or publicly-legal nature. Б.а.
39. Check Hiks, John R. "Стоимость и капитал", publisher"Прогресс", Moscow 1993г., page.102.
40. In the previous introduction was proven that this occurrence do not have a temporal nature i.e. that it is "stable" and "system-inherent"
41. Check Philip Rachev and Ilonka Goranova – "Social orders", publisher: БЦ "Информа интелект" 2005., page 8; more in this direction – Milana Krivachka and Iordan Savchev – the social orders: approaches for applying the changes in LSO ; accountance" publisher: "Personal Consult" 2002
42. Classification taken from Goleva, Lachezarov and Ruschev, "Social orders" Publisher: "Labour and law", Sofia, 2002, More and detailed on this matter – there again, pages 10-15.
43. Check Hrusanov, Kostov and Dermendjiev "Administrative Law of the Bulgarian Republic. General part" publisher "Sibi", Sofia 2001 pages 48-56
44. exactly the presence of factual composition, i.e. normatively defined (completed) precautions when realizing (arising) the relations between the subject of the public sector and the subject from the private one presents an empiric document in "the system proof" of tri-sector in the mixed economics – "to system lead only those relations of determination between the elements, that do not violate the inner structure and functions (Constitutionally determined) of the connected elements so that after the connection the elements to continue to be what thew were before their relation… as a result from their relation something new is built that is not present in each of the elements. This new thing, created as a result from the relation between the elements is the system (subsystem)" The text in italic – mine, quoted from Manov, Vasil "o Manov, Vasil "Prognosis and Planning the development and the functioning if the economic systems" – publisher UNWE – "Economics" department, Sofia, 2001 page 27
45. Precisely those elements from the factual composition, add the characteristic of determination of the relations between the publicly-legal and privately-legal subjects.

Model for Strategy of Technology New Ventures of Local Market

SiaTsolova
Sofia University "St. Kliment Ohridski", Republic of Bulgaria
E-mail: s.valentinova@gmail.com

Abstract. This report aims presenting the results from a research on modeling of strategies for technology (IT) new ventures by proposing a model for strategy for a technology new venture working on local new market. The developed model is subject of further quantitative and qualitative verification, research and approbation and is developed according a described in this report process for modeling of strategies for technology new ventures, which is used as a base for modeling of series of strategy models developed by the author. The report contains overview of the key success factors for the researched strategic model of technology new ventures working on local new market, also strategy maps of the key success factors and strategic goals in the proposed strategy model, followed by further development of the strategy model of the proposed type of technology new ventures. The proposed model for strategy for technology new ventures is subject of further analysis and approbation and it will be used as a base for further research, tests and experiments on the topic.

Keywords: strategy, model, technology, entrepreneurship, new ventures, strategic, formulation.

1. Introduction.

This report aims to present a model for strategy of technology new ventures which have a high degree of innovation capacity and start their business with products or services offering them to local emerging (new) market. The report presents the key success factors for starting an innovative technology company, the strategic business objectives of the type according to the proposed strategy, strategic maps of the key success factors and strategic objectives, showing relationships between targets of different categories and perspectives in the Balanced Scorecard from performance and definition of appropriate indicators of results to tracking progress towards each of the defined strategic objectives. The report is consistent with the strategic management process and the place of modeling strategies in it. The proposed model is based on a developed framework for the classification of strategies, and the development process and methodology for modeling strategies for start-up technology companies, the proposed model is used to further fully develop the framework and methodology of modeling strategies. The model is subject to further testing, experimentation and approbation by the author, and may serve as a basis for further research in the field of modeling strategies for start-up technology companies.

2. Process of modeling strategies and a model for strategy for technology new ventures.

Modeling strategies for technology new ventures is a matter of utmost importance for the success of any startup company. And due to the specificity of start-up companies and the inapplicability of some of the tools in the classical strategic management there is an increasingly felt need to study and analyze these specifics and construction of tools, methods

of developing and developed models to support strategic management of such companies. The proposed model is created based on the basis of proposed in this report developed process for modeling of strategies and a developed framework for classifying strategies for start-up technology companies. The process of strategy modeling begins with developing a vision of the company, followed by the identification of key factors for success in the field, which are extracted from the vision of the company. After the identification of key success factors a definition of the strategic objectives is made, derived from the key success factors, and the relationships among the key factors and strategic objectives are presented by strategic maps. Further modeling of strategy goes through establishment of indicators by which the company will monitor its progress in achieving the strategic objectives, followed by target values, actions to achieve these goals, and prioritization of certain activities in order to implement the activities under their degree of importance for achieving strategic goals. Subsequently to be responsible for the implementation of the actions and choices of deadlines is part of bringing the company's readiness to implement the defined strategy and is specific to each company, therefore not part of the modeling strategy for start-up technology companies presented in this report.

2.1. Model for strategy of technology new ventures.

This report provides a model strategy for startup technology company with high innovative capacity, market introduction of innovative products, offering products / services to local emerging (new) market.

The advent of emerging local market with innovative products from startups technology companies is one of the most interesting strategies for consideration in strategies for start-up technology companies in the local market , as it is this type of companies offering innovative products and services of local emerging market which then enter as innovative companies in the global market. That advantages of companies offering services or products for emerging local market are arising from the position of the first company to market a product on the local market, as well as technological capabilities and excellence of the company. Features in threats to such type of companies come from that – at what point in the development of the market competing companies appear, and the emergence of competing companies offering products with competitive technologies. Inovative emerging local market is characterized by innovative companies offering products to existing local market, where the limit in competition for companies operating in emerging market is the global market. This strategy has the following key features:

- In terms of product - market: Maturity of the market: emerging (new) market; Market uncertainty: medium; Potential for growth: average.
- In terms of the technological capabilities of the company: Technological capabilities: innovator; Priority research and development: quality of the product; Investments in research and development: Medium; Technology partners: local partners.
- In terms of target market: Target Market: Local; Degree of internationalization: medium.

Typical key features of the strategy: The described types of products (or services) will be offered at local emerging (new) market, offering closely innovative product or service for local emerging (new) market. Due to the high technological innovation capabilities of these companies and the fact that the boundary of competition in the market for a new product at the local level is the global market, some of these companies may also continues their growth and

development in the global market, as participants in the global market or an existing global emerging (new) market, which , however, require high levels of investment in research and development and high technological and innovative capabilities of a company. Benefits for local companies in emerging markets also come precisely from these technological innovation opportunities, but the level of investment in research and development is not as high . Feature in some companies that seek to develop international market is that the launch of innovative products for the local market is emerging as a stepping stone to improve the product and its subsequent entry to the global market. Development of innovative products for local emerging market is associated with a moderate market uncertainty, as usually similar kinds of products are sold somewhere on a foreign market, but in another country. With the availability of products that are innovative and are targeting the global market, however, this uncertainty is higher, but they will be addressed in subsequent strategy reflecting these details. The implementation strategy for innovative local companies at emerging market is related to the presence of medium level investment in research and development of new products (compared with the presence of high investments in global emerging market) and technology partners in the company's – in most cases local. At this point at innovation management in the introduction of innovations in the market for emerging (new) market the focus is on the quality of the product, as this will be crucial for the success of the product in this early life stage. The degree of internationalization of the product is average and depends on the target market and the availability of strategic development solutions to the global market in the future.

For further modeling of this type of strategy the developed process will be used for modeling of strategies, beginning with defining the company's vision and extracted from it the key success factors for achieving this vision.

Typical vision: The vision of this company is to offer an innovative product developed by the company, for the local emerging (new) market, with higher quality, seeking leadership in an area of local emerging (new) market. Such a strategy is suitable namely for start-up companies that have a high potential for innovation in the field, which also have the know - how knowledge or unfair competitive advantage, and can develop and provide the market with an innovative product or service to local emerging market. Such companies have the advantage of the first company offering a service or product on the market that rely very much on technological capabilities that are a source of competitive advantage for these companies. The focus of the company at this stage is on the quality of the product in a way that ensures the higher added value of the product or service for consumers and clients. The vision of this company determines the further key success factors associated with providing of innovative product training to customers, investments in research and development and a high level of qualification of the staff, etc.

Key to the successful implementation of this vision are: High financial results and increasing profit; Preference of customers to our product over that of its competitors; Good company image; High product quality; Increasing number of customers; Reliable operation of product; Quality work and facilitate the use of the product or service; Bearable higher price corresponding to the value added; Informing customers about the benefits of the product or service; Maintain a high quality product; Quality customer service; Displaying the innovation market; Qualified personnel; Synchronizations and teamwork; Motivating working atmosphere.

These key success factors determine the priorities of the company in order to implement this type of strategy and are interconnected due to the fact that success in some of these factors leads to success of others or support the successful achievement of others. These relationships are visualized in a strategic map of the key factors for success, built on rules for the balanced

development of the company in the following sections: finance, clients or customers, products, company internal processes and learning and growth.

A change in the classic balanced scorecard, which is developed in this report, is the existence of a strategy perspective – product perspective, describing how the product reflects customer needs and how the internal processes are leading to the successful implementation of that product or service. Therefore Perspective Product is located between Clients perspective and internal processes perspective. The developers of the balanced scorecard - R. Kaplan and D. Norman - state that depending on the different needs and future developments of companies in different sectors and the names and number of perspectives can be changed. The balanced scorecard is a tool designed for the strategic management for development companies with multiple departments and resources that need complex coordination between various departments in horizontals and also in vertical direction. Using cascading balanced scorecard maps for balanced outcome in itself is not applicable for startups because of their specificity - a small number of staff, limited resources, etc. And also main focus is not on coordination. The presented process and modeling strategy for start-up technology companies in this report uses the balanced scorecard base and also a revised modeling process, which is adapted and suitable for the strategic management of start-ups. In the classic form of balanced scorecard there is a lack of perspective products, but due to the focus of the work of startups on the modeling of the product and search for a suitable form of the product and its supply and also the needs of entrepreneurs, after providing interviews with entrepreneurs, a perspective product was introduced at the balances scorecard that exactly matches the needs of start-up companies and supports the synergy of work in the initial steps of technology new ventures among the various tools they use, and which are aimed at different aspects of the business, but always involving the company's product as main ingredient.

Scheme.1: Updated balanced scorecard model

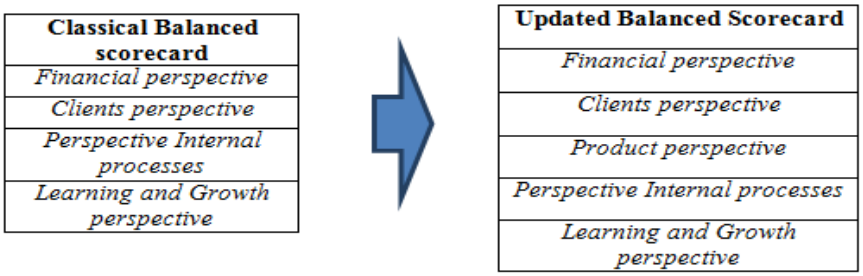


Figure 1: Strategic map of key success factors.

Strategic map of the key success factors helps to determine the most important factors for a given strategy, which reflects in further prioritizing the activities related to these factors. After the identification of critical success factors, an identifying of the main strategic goals for the company to set key success factors. For each key success factor the company defines at least one strategic goal following balanced scorecard structure, as these objectives are aimed at the further development of a set of activities leading to the achievement of these goals.

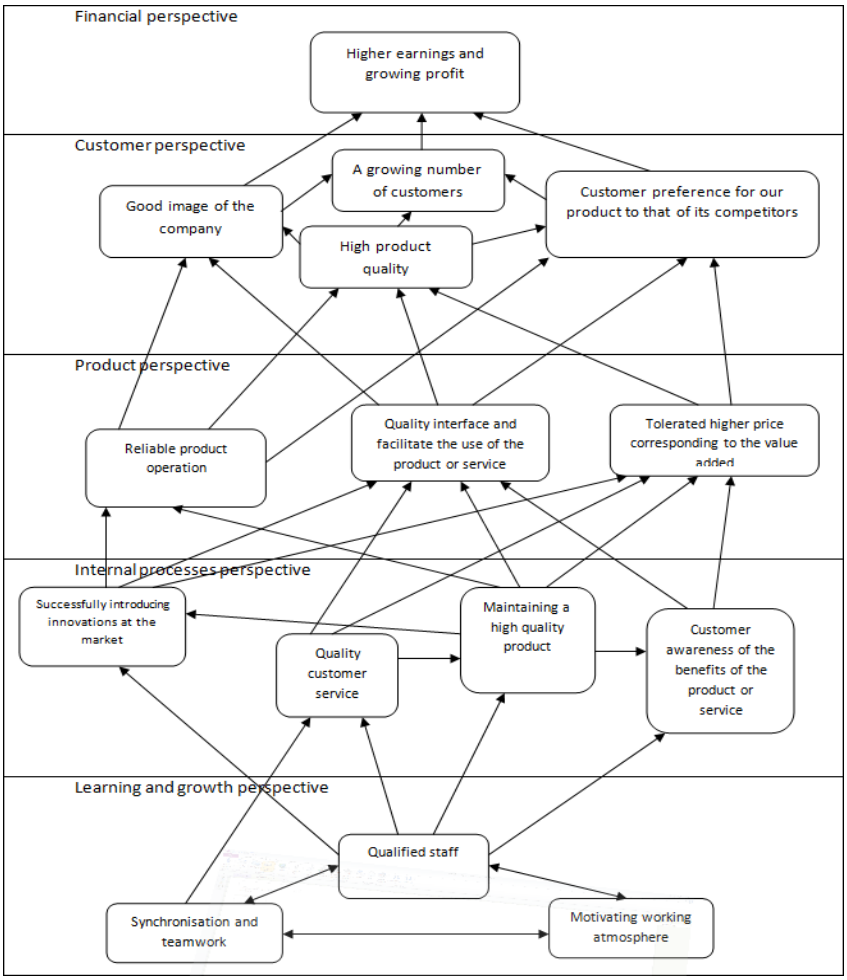


Table 1: Key success factors and related strategic objectives.

Key factors for success	Strategic Objectives
Financial perspective	
Higher earnings and growing profit	Increase the company's value
	Increasing Profit
Customer perspective	
Customer preference for our product to that of its competitors	Growing relative market share
	A growing number of customers
	Reduce the cost of acquiring new customers or retain old one
Good image of the company	Sustaining image of a reliable company offering specialized products
High product quality	Offering products of exceptional quality with innovative features

A growing number of customers	A growing number of customers
Product perspective	
Reliable product operation	Reliable product operation of the product or service
	Timely updates and product updates
	High degree of customer satisfaction
Quality interface and facilitate the use of the product or service	Create quality interface design and facilitate the use of the product or service
Tolerated higher price corresponding to the value added	Keeping bearable higher price of the product corresponding to the added value for customers
Internal processes perspective	
Customer awareness of the benefits of the product or service	Clearly inform customers about the benefits of the product or service
Maintaining a high quality product	Maintaining a high level of staff training in
	Quality control of the product development
	Quality product testing
Quality customer service	Quality service when buying
	Quality service at support
	High degree of customer satisfaction
	Timely updates and product updates
Successfully introducing innovations at the market	Continuous staff training
	Successful introduction of market innovation
	Active feedback from users to improve product
Learning and Growth perspective	
Qualified staff	Continuous staff training
Synchronisation and teamwork	Good teamwork
	Stimulate employee suggestions to improve business processes
	Development of business processes to an active exchange of ideas and experiences between staff of different directions
Motivating working atmosphere	Building a strong organizational culture
	Build trust and improve informal communication between employees
	Continuous staff training

A listing of the company's strategic goals is followed by the mapping of the interconnections of these strategic objectives by building strategic map of strategic objectives. Map of strategic goals helps to prioritize the most important to achieve strategic objectives.

Figure 2.: Strategic map of company's strategic goals. (see on next page)

The successful implementation of each activity requires setting the correct order and tracking progress in the implementation of activities leading up to the goal. To achieve the strategic objectives, key performance indicators by which to monitor progress in achieving the objectives should be set, indicators should be quantifiable. The next step in the process of modeling strategy is the definition of performance indicators for each strategic objective.

After determining the key performance indicators, the maximum level targets to achieve, for a chosen period, should be determined. They are specific to each company and are associated with a detailed analysis of the internal and external environment for the company. Further examination and monitor performance are carried out precisely on these indicators. In

order to reach the strategic goals, the company determines the main activities to implement in order to achieve the strategic objectives for each indicator according to which the company will report the result of the success of these activities by the set targets for the certain key indicators. In this step of strategy modeling, the company will have full information about the chosen strategy - from the strategic vision (the vision of the company) to the main activities that must meet a set relationships between key success factors and activities to achieve the strategic objectives of start-up technology company. Relationships between key success factors, strategic objectives, result indicators and activities to achieve the goals of modeling proposed in this report strategy are shown in the following table.

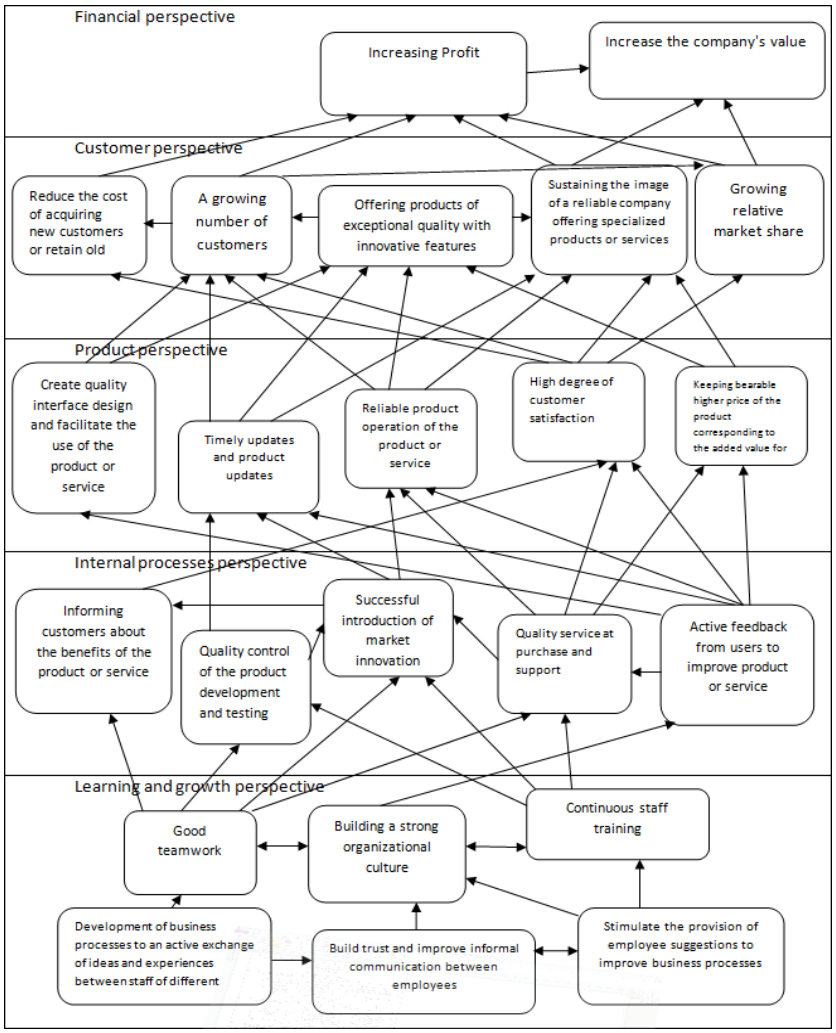


Table 2.: Relationship between key success factors, strategic objectives, result indicators and activities to achieve strategic goals.

Key factors for success	Strategic Objectives	Key performance indicators	Target values	Actions to achieve the objectives
Financial Perspective				
Higher earnings and growing profit	Increase the company's value Increasing Profit	value of the company number of sales value of net worth compared to indirect costs	* To be determined	Periodic analysis of the company's price Implementation of marketing policies to increase the number of sales Reducing the value of indirect costs
Customer perspective				
Customer preference for our product to that of its competitors	Growing relative market share A growing number of customers Reduce the cost of acquiring new customers or retain old one	% Market share number of customers number of new clients number of loyal customers The cost of acquiring a customer	* To be determined	Expanding the product range ad campaign Expanding the product range ad campaign Application of optimal pricing Optimizing marketing expenses Applying optimal pricing policy
Good image of the company	Sustaining the image of a reliable company offering specialized products or services	level of customer satisfaction number of complaints	* To be determined	Applying public relations and social responsibility Receive feedback from customers
High product quality	Offering high quality products with innovative features	degree of product innovation	* To be determined	Development and subsequent upgrade of enhancing quality innovative elements of the product or service
A growing number of customers	A growing number of customers	number of customers number of new clients number of loyal customers	* To be determined	Promotion of innovation and capabilities of the product or service ad campaign held at an appropriate price policy
Perspective Product				
Reliable product operation	Reliable product operation of the product or service Timely updates and product updates High degree of customer satisfaction	number of defects; number of com-plaints; number of canceled sales period period of implement-tation of improvements in identifying their needs number of updates; level of customer satisfaction	* To be determined	Improving the testing process of product study of the most common problems in functioning;Shortening the period to remedy the identified problems in functioning; Shortening the period to eliminate the identified problems in functioning; Shortening the period for upgrades and updates Receive assessment and feedback from customers
Quality interface and	Create quality interface design and facilitate the	Positive feedback from customers	* To be determined	Creating a Quality interface and facilitate the use of the product or

facilitate the use of the product or service	use of the product or service	level of customer satisfaction		service Receive assessment and feedback from customers
Tolerated higher price correspondi ng to the value added	Keeping bearable higher price of the product corres-ponding to the added value for customers	position in the price category	* To be determined	Study of competitive pricing and products Informing customers about the benefits of the product or service
Internal process perspective				
Customer awareness of the benefits of the product or service	Informing customers about the benefits of the product or service	volume of information provided positive feedback from customers	* To be determined	presenting information to customers about the benefits of the product or service Research on the level of understanding by customers of the benefits of the product or service
Maintainin g a high quality product	Quality control of the product development Quality product testing	number of product quality inspections number of products passing the quality check; number platforms for which the product is tested; number of test modules of the product	* To be determined	Develop a strategy for checking the quality of the product and the process of its implementation Develop a strategy for testing the quality of the product
Quality customer service	Quality service when buying Quality service at support	level of customer satisfaction; number of complaints; level of satisfaction of customers; number of complaints; avg. number of calls from customers at issue	* To be determined	Process optimization in the purchase of the product Training of staff Feedback from customers Process optimization of product support training of staff Feedback from customers
Successfull y introducing innovations at the market	Successful introduction of market innovation	number of successfully drive innovation in the market as part of a product or service of new products or services	* To be determined	Study the needs of the market and consumers Develop a policy for decommissioning strategies Staff Training Analysis of the deduced innovation
	Active feedback from users to improve product	number of contributions from consumers; number of users giving feedback	* To be determined	Optimisation of the process of preparation of the product ideas Feedback from customers
Learning and Growth perspective				

Qualified staff	Continuous staff training	number of trainings;% Staff trained; % Employees passed the exams; % Employees qualify for the level of qualification	* To be determined	Preparation of a staff training Training of staff Analysis of the results of the training
Synchronisation and teamwork	Good teamwork	number of regular meetings of rep. from different directions; No of team-buildings	* To be determined	Conducting regular meetings of officials from different directions Conducting team-buildings
	Stimulate the provision of employee suggestions to improve business processes	No of team-buildings; number of proposals to improve business processes	* To be determined	Create an environment to give suggestions to improve the business processes of employees; An active policy to stimulate the provision of employee suggestions Conducting team-buildings
	Organise env. for active exchange of ideas and experiences between staff of different sections	number of regular meetings of officials from different directions No of team-buildings	* To be determined	Building processes and systems stimulate the exchange of ideas and experiences between staff of different directions; Conducting regular meetings of officials from different directions; Conducting team-buildings
Motivating working atmosphere	Building a strong organizational culture	No of team-buildings; number of company-branded products in company; number of regular meetings of reps. from different directions	* To be determined	Creating and implementing a strategy for building a strong organizational culture
	Build trust and improve informal communication between employees	number of regular meetings of officials from different directions No of team-buildings	* To be determined	Conducting regular meetings of officials from different directions Conducting team-buildings Building a center of communication in the workplace
	Continuous staff training	number of trainings % Staff trained % Employees passed the exams	* To be determined	Preparation of a staff training Training of staff Analysis of the results of the training

After building the basic framework of information the next step is proceeding to ordering activities by importance, which in turn is related to the prioritization of each of the activities against the strategic importance of key factors whose success it supports.

Table 3: Priority actions for achieving the objectives.

Key factors for success	**Actions to achieve the objectives**	**weight of KFS (max-4 min-1)**	**weight of operation (up to 4-min-1)**	**Priority (up to 16-min-1)**
Financial Perspective				
Higher earnings and growing profit	Periodic analysis of the company's price	4	2	8
Higher earnings and growing profit	Implementation of marketing policies to increase the number of sales	4	4	16
Higher earnings and growing profit	Reducing the value of indirect costs	4	3	12
Customer perspective				
Customer preference for our product to that of its competitors	Expansion of the product range	4	2	8
Customer preference for our product to that of its competitors	Advertising campaign	4	4	16
Customer preference for our product to that of its competitors	Application of optimal pricing	4	4	16
Customer preference for our product to that of its competitors	Optimizing marketing costs	4	3	12
Good image of the company	Applying public relations and social responsibility	4	2	8
Good image of the company	Feedback from customers	4	4	16
High product quality	Development and subsequent upgrade of enhancing quality innovative elements of the product or service	4	4	16
A growing number of customers	Promotion of innovation and capabilities of the product or service	4	4	16
A growing number of customers	Advertising campaign	4	4	16
A growing number of customers	Conducting appropriate pricing policy	4	4	16
Product perspective				

Reliable product operation	Improving the testing process of product	4	4	16
Reliable product operation	Study of the most common problems in functioning	4	4	16
Reliable product operation	Shortening the period to eliminate the identified problems in functioning	4	4	16
Reliable product operation	Shortening the period for upgrades and updates	4	3	12
Reliable product operation	Receive assessment and feedback from customers	4	4	16
Quality interface and facilitate the use of the product or service	Creating a Quality interface and facilitate the use of the product or service	4	4	16
Quality interface and facilitate the use of the product or service	Receive assessment and feedback from customers	4	4	16
Tolerated higher price corresponding to the value added	Study of competitive pricing and products	3	4	12
Tolerated higher price corresponding to the value added	Informing customers about the benefits of the product or service	3	4	12
Internal process perspective				
Customer awareness of the benefits of the product or service	Presentation of information to customers about the benefits of the product or service	4	4	16
Customer awareness of the benefits of the product or service	Research on the level of understanding by customers of the benefits of the product or service	4	4	16
Maintaining a high quality product	Develop a strategy for checking the quality of the product and the process of its implementation	4	4	16
Maintaining a high quality product	Develop a strategy for testing the quality of the product	4	4	16
Quality customer service	Process optimization in the purchase of the product	3	3	9
Quality customer service	Training of staff	3	4	12

Quality customer service	Feedback from customers	3	4	12
Quality customer service	Shortening the period to eliminate the identified problems in functioning	3	4	12
Quality customer service	Shortening the period for upgrades and updates	3	3	9
Successfully introducing innovations at market	Preparation of a staff training	4	4	16
Successfully introducing innovations at market	Training of staff	4	4	16
Successfully introducing innovations at market	Analysis of the results of the training	4	3	12
Successfully introducing innovations at market	Study the needs of the market and consumers	4	4	16
Successfully introducing innovations at market	Training of staff	4	4	16
Successfully introducing innovations at market	Analysis of the deduced innovation	4	3	12
Successfully introducing innovations at market	Optimisation of the process of preparation of the product ideas	4	4	16
Successfully introducing innovations at market	Feedback from customers	4	4	16
Training and Growth perspective				
Qualified staff	Preparation of a staff training	4	4	16
Qualified staff	Training of staff	4	4	16
Qualified staff	Analysis of the results of the training	4	4	16
Synchronisation and teamwork	Conducting team-buildings	3	3	9
Synchronisation and teamwork	Create an environment to give suggestions to improve the business processes of employees	3	3	9
Synchronisation and teamwork	An active policy to stimulate the provision of employee suggestions	3	3	9
Synchronisation and teamwork	Building processes and systems stimulate the exchange of ideas and experiences between staff of different directions	3	3	9

Synchronisation and teamwork	Conducting regular meetings of officials from different directions	3	4	12
Synchronisation and teamwork	Conducting team-buildings	3	3	9
Motivating working atmosphere	Creating and implementing a strategy for building a strong organizational culture	4	4	16
Motivating working atmosphere	Building a center of communication in the workplace	4	4	16
Motivating working atmosphere	Conducting team-buildings	4	3	12
Motivating working atmosphere	Conducting regular meetings of officials from different directions	4	4	16
Motivating working atmosphere	Preparation of a staff training	4	4	16
Motivating working atmosphere	Training of staff	4	4	16
Motivating working atmosphere	Analysis of the results of the training	4	4	16

Setting priorities in the implementation of activities is followed by the final step to determine those responsible for the implementation of their respective activities. This table contains basic information which is enriched with information about the priority of each task and responsibility, which in turn is a smooth transition to the next step in the strategic management - namely the implementation of the strategy and its execution. After implementation of these steps the company has outlined a strategy and is ready for its implementation.

The model for the strategy of starting a technology company covers local companies with highly innovative operating at local emerging (new) market and is the subject of further research and experiments.

This report presents application of the developed process of strategies modeling for technology new ventures in modeling of typological strategies for technology start-up companies by presenting a developed model for technology new venture with high innovative capacity, offering products at local emerging (new) market. The process and also the model are subjects of further research and experiments.

3. Conclusion

Modeling strategies for start-up technology companies is characterized by many features and factors of external and internal environment for the companies. The rapid development of technological entrepreneurship requires study of the major types of strategies for start-up companies develop appropriate processes and tools for modeling strategies, identifying key success factors, etc. The proposed model for strategy of technology new company covers tech start-ups with a high degree of innovation opportunities, operating on local emerging (new) market and offering products or services that are not available at that time of the chosen market. The model contains the main factors and elements of the developed

process for strategy modeling of start-up technology companies, which process and model are subject to further development, qualitative and quantitative research and experiments in the field of modeling of strategies for technology new ventures.

References

1. Harvard Business Press (2005), "Strategic management, Harvard Business Essentials".
2. M. Panova and T. Panov (2008), "Strategic management".
3. R. Kaplan and D. Norton (1996), "The Balanced Scorecard: Translating Strategy into Action, Harvard Business School Press".
4. Steven Gary Blank (2006), "The Four Steps to The Epiphany – Successful Strategies to Products that Win".
5. Jeffry Timmons (2006), "New Venture Creation, Entrepreneurship in the 21st Century".
6. Steve Blank and Bob Dorf (2012), "The Startup Owners Manual, The Step by Step Guide to Building a Great Company".
7. Petko Ruskov (2009), "Strategic Model and Framework for IT/IS GraduateEducation".
8. Petko Ruskov, Mark Harris (2007), "Strategic Model for Master of Science Program "Innovation and Technology Entrepreneurship".

Standardization of Public Sector Accounting in the EU

Daniela Feschiyan,
Department of "Accounting and Analysis", UNWE, Sofia

Abstract: This paper analyzes the need for harmonizing the standards for accounting reporting by fully applying the principle of Accrual Based Accounting in the public sector in the EU member states. It also examines the discrepancy between the accountancy in the public sector, based on the cash flow principle (cash basis) and the Accrual Based principle applied in the process of budgetary surveillance in the EU (ESA 95, 2011). The emphasis is laid on the strong dependency of the statistical information at the European level on the quality of the accounting and information basis of public finances and on Eurostat's willingness to create a system of harmonized accounting standards, based on the Accrual Based accounting, which are in compliance with ESA and applicable to all agents in the State Management sector. This paper further analyzes the standardization of the accounting in the public sector of the EU member states and globally and presents some trends of its development. The conclusion is reached that IPSAS are an indisputable staring point to develop harmonized reports in the EU public sector. IPSAS provide the best basic framework for the development of a set of European Public Sector Accounting Standards.

Keywords: accounting, public sector, Accrual Based Accounting, European Public Sector Accounting Standards

Introduction

The discrepancy between the information reported by the public sector and the data calculated on the accual basis in compliance with ESA 95 standard is highlighted in the communication released by the European Commission to European Parliament and the Council on 15 April 2011 "Towards robust quality management of European statistics". This communication focuses on the strong dependency of statistsical information at the European level on the quality of the accounting and information basis of public finances. This is the reason why Eurostat insisted on establishing a system of harmonized accounting standards based on the principle of accrual accounting in line with ESA for all agents functioning in the State Management sector.

In 2012 the European Commission made an assessment of the application of the International Accounting Standards in the public sector (International Federation of Accountants, 2012) within the European Union (EC, 2012). This assessment triggered a broad debate on the problems related to the standardization of the accounting system in the public sector. An analysis of the findings of the public survey conducted by the Eurostat Directorate General with the participation of the Federation of European Accountants (FEE) and the conclusions it presented prove the need that financial statements in the public sector should be harmonized in compliance with the principles of accrual-based accounting. Hence there is an objective need to develop European accounting standards for the public sector (EPSAS) based on the rules and provisions and the overall rationale of the International Public Sector Accounting Standards (IPSAS).

The International Public Sector Accounting Standards

The major goal of the commission for the public sector is to develop programs intended to improve the financial management and accounting in the public sector by:

- ✓ The development of accounting and auditing standards and promoting their adoption;
- ✓ The development and coordination of all programs to promote training and research;
- ✓ The promotion and facilitation of the information exchange between the member organizations of the Board at the International Federation of accountants and other stakeholders.

The International Public Sector Accounting Standards are not regarded as having priority over the national provisions in legislation. IPSAS reaffirm the full discretion of governments and national departments in their choice of a model for the organization of the accounting systems between applying IPSAS and applying the national accounting standards for the public sector. With a view to improving the quality of the financial statements in the public sector worldwide it is strongly recommended that IPSAS be adopted and applied. The author is a proponent of the theory that IPSAS should be fully applied, which will result in raising the quality of financial statements and improving the transparency in the public sector. We also adopt the approach that most European countries apply – the development of national accounting standards for the public sector based on the principles and premises of IPSAS (Feschiyan, 2010, p.132) under the strict observance of the principles of accrual-based accounting (Stoyanov, Feschiyan, 2012, p. 71). IPSAS are applied by all agents in the public sector – national governments, regional units of local government, local authorities, their accompanying departments and other agents, excluding the state-owned economic enterprises. As was mentioned above, they are developed and issued by the the commission for the public sector which is an integral part of the International Federation of Accountants Council (IFAC) seated in New York, the USA. IPSAS were developed on the basis of the International Accounting Standards. Their complete contents are posted on the official website of the International Federation of Accountants both in English and many other languages such as French, Spanish, German, Italian, Russian and Chinese. This is the only set of accounting standards in the public sector that enjoys international recognition. They are based on the assumption that the modern management of the public sector in compliance with the principles of cost efficiency, effectiveness and efficiency is dependent on the managerial information systems that provide the timely, precise and reliable information about the financial and economic health of the country and the performance of the respective governmental agencies or another economic agents.

The International Public Sector Accounting Standards are aimed at creating comparable, relevant and legible financial information to meet the needs of both internal management and international institutions and rating agencies, capital markets and other institutions. They provide instructions and guidelines pertaining to the structure of the financial statements, including the minimal requirements for each basic statement, accounting policy and the explanations (and the appendices which serve as their illustrations). They idea is that the adoption of the International Public Sector Accounting Standards should provide the basis for the enterprise to present its general-purpose financial statements in such a manner that they be comparable with the statements of previous reporting periods of the enterprise itself and with those presented by other enterprises outside the public sector. It is namely through the application of the accounting standards that the general lines of presenting the

financial statements are given as well as guidelines for their structure and the minimal requirements with regard to content. As of the present moment 33 IPSAS have been approved and applied, of which 32 standards follow the principle of accrual-based accounting and one standard follows the cash flow base principle. These are the major bases on which public sector enterprises should prepare their financial statements. In preparing the financial statements in compliance with accrual-based accounting, the transactions and the events should be recongnized at the moment of their arising, regardless of the moment of payment or the receipt of money or monetary equivalents. In preparing the financial statements in compliance with cash flow base principle, the transactions and the events should be recongnized at the moment of their payment or the receipt of money. A basic aspect of the rationale behind IPSAS is that public sector enterprises should apply the accrual-based accounting principle. At the heart of this principle is the idea of the implementation of the comparability concept. It is related to the recognition of revenues and expenditures and requires that the latter be recognized in the accounting documents at the moment they arise, not at the moment of their payment or the receipt of funds. Hence the balance sheet reflects the arisen accounts receivable and payable, whereas the income statement presents both the monetary and non- monetary revenues and expenditures. In our view the application of the accrual-based accounting principle in the public sector ensures that the following positive effects be achieved:

- ✓ Creating the conditions for and taking on the commitment to carry out a systemic surveillance and assessment of the collection rate of receivables;
- ✓ Providing the basis for the implementation of the realization concept, which is related to the recognition of revenues in accounting;
- ✓ Improving the contents of the financial statements – apart from the information of previous transactions related to payment or the receipt of monetary funds, they should also contain information about future accounts payable or receivable, which at some future moment will appear as items of incoming or outgoing cash flows;
- ✓ Presenting all resources that are subject to the control of the government;
- ✓ Creating the information basis to make an analysis of the cost-effective and efficient functioning of the enterprise;
- ✓ Creating a good information basis to make managerial decisions related to the enterprise's planning and control of its activities;
- ✓ Revealing the effect of past decisions on the current financial state as well as the the effect of current decisions on future financial state;
- ✓ Revealing the efficiency in the utilization of assets (Feschiyan, 2013, p. 402) ;
- ✓ Providing the information basis to improve the management of accounts receivable;
- ✓ Creating information about the net assets, which will provide for improving the management of public debt;
- ✓ Creating conditions for public sector enterprises to present their assets and performance in a true and fair manner (Feschiyan, Raleva, 2013, p. 388).

In structural terms the International Public Sector Accounting Standards have strong resemblance to the International Financial Reporting Standards (IFRS). However, in terms of content there are differences that stem from the specific features of public sector enterprises. These differences are determined by the specific aspects of the activities of these enterprises, namely:

Lack of financial indicators oriented to profit;

The following are the three basic indicators of the assessment of the performance in place:

- Cost efficiency – achieving the goals with the least resources;
- Efficiency – achieving the goals with the most optimal utilization of the resources;
- Effectiveness – achieving the ultimate goal.

To meet the growing global demands the standards should be open to changes, further development and improvement (Feschiyan, 2010, p.1557).

Harmonization of accounting in the public sector in EU

The trend in the standardization of the EU member states' public sector accounting is toward the adoption and application of harmonized European accounting standards, based on the principles, rules and provisions of IPSAS in full compliance with the accrual-based accounting principles. The expectations are that the adoption of a single set of accounting standards that follow the the accrual-based accounting principles at all levels of State Management in the EU will produce significant effects for public sector management. The reporting in compliance with the accrual-based accounting principles in public sector enterprises will boost the effectiveness and efficiency of public administration and facilitate the ensuring of liquidity, which is necessary for the maintenance of viable public services.
The double-entry reporting in compliance with the accrual-based accounting principles provides a reliable accounting framework that can provide the necessary information in a feasible and timely manner. Public sector enterprises' compliance with the accrual-based accounting principles will improve transparency, management responsibility and the comparability of financial reporting. This will inevitably result in increased effectiveness and efficiency of public auditing.

The presented data show that at present the national accounting standards in the public sector in most EU member states are not directly related to IPSAS. Most of the 15 EU member states have based their national standards on IPSAS, or have brought them in line with IPSAS, or have applied IPSAS directly in some parts of local government or at least draw upon IPSAS provisions. Even though IPSAS value has been widely recognized, however, not a single member state has yet fully applied them (European Commission, 2013, p 9).

In 2012 the framework for IPSAS management was subject to revision (European Commission, 2013, p 5) with a view to eliminating the problems that the applying authorities and stakeholders had spotted. Most EU member states united around the idea of developing a set of European public sector accounting standards (EPSAS). The EU member states applying IPSAS share the view that they would be appropriate as a referential framework for the development of EPSAS in the future. The idea is that these standards will provide the EU with the opportunity to prepare their own standards that meet their own requirements at the necessary pace. They are envisaged to contain a set of harmonized accounting standards following the accrual-based accounting principles in the public sector, adapted to the specific needs of the EU member states. The application of EPSAS throughout the EU will considerably reduce the complexity of the preparation methods and procedures which are used in the transformation of these data on a quasi-harmonized basis and will curb the risk with regard to the reliability of the information that the EU member states have notified of and which have been published by Eurostat (European Commission, 2013, p7).

The first step in this transition must be the introduction of this project management at the EU level in order to clarify the conceptual framework and the application of a common financial reporting in the EU public sector. The harmonized EPSAS should be based on the

principles, rules and provisions of IPSAS. We deem it appropriate that the IPSAS that are generally accepted by the member states should be directly integrated within the harmonized EPSAS framework. At the same time there should be no restrictions in the development of the new standards. Based on the experience of the countries that have applied the accrual-based accounting principles in the public sector in the past years, the process of adopting single IPSAS-based European standards is envisaged to be carried out stage by stage. As the first stage the initial situation in each member state should be analyzed, which includes the state of the national accounting standards and the existence of an adequate legal framework. Some member states may find it appropriate to launch their implementation at the national level, and at a later stage — at the regional and local level. It should also be expected that they will be implemented at a restricted level by the smaller entities, or that at least the more relevant entites should be given priority in this respect, considering their significance.

This process is to to be carried out in three stages:

Preparatory stage - the collection of more information and views and the preparation of a pilot project. This stage may be launched as early as 2013 and include additional consultations, top level conference and the preparation of more detailed proposals.

Intermediate stage – stage at which the practical agreements are drafted and implemented, including financing, management, possible useful interaction and the problems of the smaller agents in the public sector, among others. This stage should be completed with the publication of a draft framework regulation. The framework regulation requires that the the accrual-based accounting principles be observed and outlines the plans for an additional preparation of specific accounting standards based on the IPSAS rationale.

Final stage – The application of EPSAS is launched. The process of their implementation should be incremental and provide for more time in the cases in which the existent accounting standards in the EU member states are at a serious variance with IPSAS. As a result of the lack of standardized accounting systems in some EU member states, the application is expected to be completed in the medium term.

One should take into consideration the fact that the preparation of harmonized European public sector accounting standards cannot by itself ensure the provision of timely and high-quality accounting data for the public sector. A number of additional conditions should be met to guarantee the latter, including the following, among others:

- Stable political support and correct implementation of the project;
- The readiness of all public administrations and agents in the public sector to work with a more complicated accounting system;
- Integrated information systems for the budget, the payments, contract management, double-entry accounting, the management of payments and statistical accountancy;
- Timely reporting (for instance on a monthly basis) of all economic events within the intergrated accounting system for the public sector agents;
- The availability of human resources and modern information technologies;
- Effective internal control and external financial audit of reporting in the public sector.

For all EU member states and particularly for those which are using only cash flow reporting, the implementation of EPSAS following the accrual-based accounting principles suggests that a significant reform should be conducted. In our opinion some of the problems that may possibly arise and which will have a direct impact on are as follows:

- Conceptual and technical accounting problems;
- Expertise of the employees and the consultants and training skills;

- Communication with the managers and those in charge of decision making, as well as their training;
- Connections with auditors, as well as their training;
- Correction and updating of the IT systems;
- Adaptation of the existent national legal frameworks.

Conclusion

In conclusion based on the presented overview of the available IPSAS we are expressing our categorical view about the advantages of the implementation of these standards, which provide the basis for the harmonization of accounting in the public sector of the EU member states. It is an indisputable fact that the financial statements prepared in compliance with IPSAS are high-quality ones ensuring reliability, comparability, transparency and publicity. This increases the authority of the countries that are applying IPSAS or have adopted their principles and rules in drafting the NAS. It further increases the confidence on the part of taxpayers and the international governmental and non-governmental organizations. Furthermore it improves the information basis of the financial management of the public sector enterprises, which thus comes closer to the requirements of the public and economic sector.

In 2012 the management framework for IPSAS of the EU member states was subject to revision in order to have the existent problems resolved. This revision was aimed at guaranteeing an independent procedure for determining the standards which enhances and meets the specific needs of the public sector. A report, accompanied by a working document of the European Commission services summarizing the current implementation of IPSAS in the EU member states with regard to reporting in the public sector (European Commission, 2012, p 9) and the results from the public consultation held in the period February 2012 by Eurostat in order that views be collected on the applicability of IPSAS (EU, 2013, p 3), points out some of the following drawbacks:

- The applied accounting practices are not described precisely enough, considering that some provide for making a choice between several types of accounting which in effect restrict harmonization;
- The set of standards are inadequate both in terms of scope and applicability to deal with some relevant governmental cash flows, such as tax and social payments. Furthermore, it fails to sufficiently account for the specific needs, characteristic features and reporting in the public sector. A major issue at stake is that IPSAS capacity should resolve the problem of the consolidation of the accounts in compliance with the definition of the State Management sector, which is currently a basic concept of EU's fiscal surveillance;
- Some standards cannot be deemed sufficiently stable, as they are expected to be updated after the project that aims to complete the conceptual framework ends, which is expected to happen in 2014;
- Another weakeness in IPSAS management is the insufficient participation of the European authorities for reporting in the public sector.

Despite the presented critical comments by the authorities in the EU member states and other agents in public consultation, the general conclusions drawn straightforwardly express the opinion that IPSAS are an indisputable starting point from which the harmonized reports in the EU public sector can be made.

In conclusion we are expresiing our view that IPSAS provide the best basic framework for the development of a set of European accounting standards in the public sector.

References

1. Stoyanov, S., D. Feschiyan,(2012) *Problems and Trends in the Development of the Accounting System of the Public Sector Enterprises in Bulgaria's Integration in the EU*, Scientific Studies UNWE, p 71
2. Feschiyan, D., (2010) *Adapting the Accounting of Public Sector Enterprises to the EPSAS Conceptual Framework*, Paper presented at the Jubilee international scientific and practical conference held at the Finance and Accounting Department at UNWE, Stopanstvo printing house, Sofia, p 132
3. Feschiyan, D.,(2012) *Divergence from the fundamental accounting principles in the accounting system of public sector enterprises*, Paper presented at an international conference held at the Finance and Accounting Department at UNWE, Sofia, October , volume 1, p 169
4. Feschiyan, D.,(2010) *International Public Sector Accounting Standards in Bulgaria*, Paper presented at 2nd Balkans and Middle East Countries Conference on Auditing and Accounting History, Conference Proceedings, Vol. II p. 1557
5. Feschiyan, D., (2013), *A System of Indicators for Financial Analysis of the Municipal Real Property,* Acta Universitatis Danubius. Economica, Volune 9, Issue № 4/2013, p. 402
6. Feschiyan, D. S. Raleva, (2013), *Valuation Bases and Accounting System Entries in Financial Analysis*, Acta Universitatis Danubius. Economica, Volune 9, Issue № 4/2013, p. 388
7. International Federation of Accountants, 2012, Handbook of International Public Sector Accounting Pronouncements, UK
8. Regulation (EU, 1996) № 2223/96 of the Council of 25 June 1996 on the European System of national and regional accounts within the Community (OBL 310, p 1)
9. Directive (EU, 2011) 2011/85/EU of the Council of 8 November 2011 on the requirements for the budget frameworks of the member states (OB L 306, 23.11.2011 г, p. 41)
10. Regulation (EU, 2011) № 1173/2011 of European Parliament and Council of 16 November 2011 on the effective enforcement of budgetary surveillance in the euro area, OBL 306, с. 21
11. Regulation (EU, 2011) № 1174/2011 of European Parliament and Council of 16 November 2011 on the enforcement measures to correct of the excessive macroeconomic imbalances in the euro area, (OB L 306, p 8)
12. Regulation (EU, 2011) № 1175/2011 of European Parliament and Council of 16 November 2011 amending Regulation (EC) № 1466/97 on the strengthening of the surveillance of budgetary positions and the surveillance and coordination of economic policies (OB L 306, p 12)
13. Regulation (EU, 2011) № 1176/2011 of European Parliament and Council of 16 November 2011 on the prevention and correction of macroeconomic imbalances (OB L 306, p 25)
14. Regulation (EC, 2011) № 1177/2011 of the Council of 8 November 2011 amending Regulation (EC) № 1467/97 on speeding up and clarifying the implementation of the excessive deficit procedure (OB L 306, 7.7.2005, p 33)
15. Working document (EC, 2012) of the services at the Commission, accompanied by the Report of the Commission submitted to the Council and European Parliament: Towards implementing harmonized public sector accounting standards in Member States, Brussels, p. 9.

Communication Technology in the Field of Building Management Systems

Dragomir Stefanov

UNWE, 1700 Sofia, Bulgaria
dr_stefanov@abv.bg

Abstract: "Intelligent Building " is a multidimensional concept. This is not only a building, but also a comlex of systems and technologies - a great combination of engineering solutions aiming a creation of effective, economical and meeting residents'needs, building management system. To achieve the above mentioned integration into a single infrastructure is necessary a transmission and interrelation of the basic systems into integrated data transmission environment. In order to reach optimal results, it is necessary meticulous attention to design and implementation of each component of the intelligent building. Overall building automation involves many factors - from the appropriate manageming of lighting depending on weather to modern systems for heating, ventilation and air conditioning (HVAC). In order to facilitate communication between different systems is used a single basic standardized communications protocol. Its purpose is to introduce a unified communicating relation from the management systems to the implemented elements. This simplifies the installating, maintenaning and future upgrading of the system.

Keywords: intelligent building, communications protocol, IP network.

At the beginning of the XXI century construction industry is influenced by many different factors, but two of them have more sufficient impact – Information Technology (IT) and Intelligent Buildings (IB). The first concept summarizes resources (computers, modems, servers, controllers, communication networks, various types of software, etc.) through which information is made available to people.

The second term has no strict definition. In general, the main goal of " intelligent building" is to achieve a healthy and comfortable lifestyle together with the most efficient resource usage. For the first time, the term appears in the early 80s of last century, in which period it is used to advertise building with a fire alarm system and access control. The concept is in contunues development, as result of the rapid innovations in the technology from the last couple of years. Today, "intelligent" buildings apply a set of technologies that allow them to work autonomously and to react on the preferences and needs of residents without prior programming of the desired activities[1]. Thus, the users could have effective, efficient, secure and using friendly home atmosphere resulting of the specfic construction of the gadget. Further, the useful automatic control of all its parts will allow the residents to enjoy the pleasure of easly reachible navigation.

Typical characteristics of intelligent buildings are:

- Ability to control major systems in the building. The operation of these systems has a direct impact on safety, efficiency and comfort of the users. These include:
  - ❖ systems for environmental control (heating, ventilation and air conditioning (HVAC), in each room according to the users preferences;
  - ❖ control of the premises for the penetration of started fire and fire alarm;
  - ❖ controlled access to individual rooms;

- ❖ measuring systems for individual consumption of water, heat and energy, ect. Automatic collection of reported data, without the need to provide access to premises;
- ❖ realization of functions to save energy through lighting, heating, air conditioning and ventilation control;

- Ability to manage complex systems – in accordance with a previous entered mode of operation and in accordance with the settings provided for entire specific conditions;
- Ability to coordinate actions by different systems;
- Opportunities for communication and accountability. Systems in intelligent buildings are enable to output detailed reports that can be customized according to the needs of information. They use modern technology for communication connection (eg. xml over IP), which enables automated sending messages to managers of buildings, and communication with other systems.[2]

Inherently, "intelligence" of the building is provided by constructing a system for building management. A typical example of such system is the network platform type BMS (Building Management System). The core of these systems for field automatization are the applied facilities and equipment of air conditioning systems , heating and cooling, access to premises, lighting and ventilation systems. On the base of them is designed the specific building automation system during the design and perform stage of the building construction. At these levels of the system are regulators, actuators, controllers, massive data transfer and more.[3]

BMS systems are specifically focused on management applications with a wide range of operation, as high reliability, easy installation and maintenance. That makes them very easily to use for all public buildings. To provide "intelligence" of the building, system is required to have an ability to continuously collect whole information coming from all the sensors and to be able to analyze it.

Modern building management systems are using sensors and control devices that are connected to a central control system. Some years ago, each system was connected to a control center with separate cables. However, in nowadays the new buildings are a single network implement that is used by all systems simultaneously, which helps for easy applicaton.

IP-based networks are becoming the most chosen solution to ensure communication. Most producers of building automation systems (BAS), sensors and control devices design their products in such a way, that they apply effectively in IP networks. Some of the advantages of using a single network are:

- Cost-effectiveness – many different systems (Building Automation Systems - BAS, corporate LAN, VoIP and even video signal) can operate in a single network. It is necessary the usage of only one set of cables instead of several. By technology Power over Internet (PoI) network cable, it will be possible to supply electricity to different devices, which will result in significant deacrease of the need of power safer cables;
- More capacity – IP systems communicate with a speed of 10 MB / sec in older cable systems, and on newer – 100 MB / sec. As compared to earlier technologies, the transmission rate of data is hundreds of times greater. In practice, this means a greater capacity allowing multiple systems to operate with one network;
- More flexibility – newer IP based networks can easily upgrade in order to offer higher speeds – often without the need of replacement of the cable systems. Also, into the existing network can be easily included new systems;
- Many choices – for the consutruction of the network, it can be used a variety of different products from worldwide suppliers.[2]

A major disadvantage for the introduction and the usage of a building automation system based on a unified IP network is the type of cables installed in the building. 70% of downtime of information networks is due to gaps in cable systems. Their proper organization is one of the key tasks in creating an intelligent building, as determined reliability of many other components. This specific task can creates obstacles in the worldwide application of the system.

Structured Cabling Systems (SCS), which combines multiple subsystems with different functions – local computer and telephone networks, security systems, video surveillance, etc., undoubtedly belongs to the foundation of the intelligent building. SCS has the building`s longest life cycle besides from the building structure components. Newer systems generate intensive data traffic related to the transmission of video and VoIP, which can not be secured by cables CAT5 or CAT5E. For this reason, it is recommended installation of cables CAT7, which are suitable for current implementation, but also for any future decisions. The value of this type of wiring is marginally higher, but the ensured long-term benefits are substantial.

However, structured cabling systems are still the most underrated element of the overall IT solution. Companies invest only 7% of their total budget for information and communication technologies in SCS, although they have the longest life. According to experts, professionally installed SCS based on high quality components can be costly at the initial point, but in the next few years ,it will repay its money in the form of power saving. Furthermore, the warranty of such SCS is approximatelly15-25 years, which implies for one secure and profitables investment.

Along with all the advantages of a single building network, underlying all the management systems of intelligent building, there are certain risks of contingencies related to the decline/failure in IP network or undefined interference. Possibles solutions for covering these risks can be further development in the backup systems and accessories, as well as an agreement with an alternative service provider.

There are created devices for the construction of BAS and BMS systems that communicate over a wireless link using the 802.11 standard. This type of devices limits the necessity of usage of network cables and their integral connection. This is one of the major advantage, which enables the installing of the system in hard to reachable places such as roofs, mechanical rooms, etc. At the same time, wireless communications has a number of difficulties and challenges, such as:

- signal interruptions caused by certain materials – steel, wood , concrete and almost every building material absorbs or reflects some of the signals emitted by wireless devices. This can result to worsen or eather makes the signal an unusable;
- interruptionsdue to sources of electromagnetic field – radio, fluorescent lights, electrical equipment and cables can worsen the proper transmission of the wireless signal;
- too few or too many wireless access points – in order to gain a full usage of the wireless network functions, it will be necessary careful design and proper determination of the number and location of access points.

The presence of any of these factors can lead to a reduction in signal quality, unwanted interruptions and delays in communication. In order to prevent incorrect operation of the wireless units is recommended to make a preliminary analysis of the construction areas for the presence of the above factors as well as their specific effects. In the process of the interpretaion and determination of the undesired results by the performed analysis, it may result to creation of various technical solutions that involve the installation of additional wireless repeaters, in order to improve the strength and quality of the signal and the speed of the link.

Dissemination of information between sensors and control devices of the building automation system is performed by signals on communication networks through appropriate communication protocols. In practice, these protocols are the "language" used by system to exchange data. The IT network supports this exchange as a "phone line", which takes place in "conversation" between the devices. The information is sent in small segments, called packets, and after each packet sending system receives a confirmation. To have the anticipated effect of the system, the transmitted information should be written in accordance with a specified communication protocol that is supported by the sending and receiving devices. This process is done in a half second.

In the early stage of development of BMS and BAS systems, other technology providers use their own communication protocols. This enforces the building owners to use basic products (sensors, pumps, etc.) and information infrastructure by a single provider, as devices from different manufacturers are incompatible. By the further development of the systems and their widespread practice, providers are beginning to offer an open source solutions for building automation systems, which allows users to integrate different components and products. Ones of the most popular the open source communication protocols are BACnet, LonWorks, EIB and KNX.

BACnet is a protocol for building automatic systems and control networks compatible with a wide range of all sizes of physical networks, making it suitable for use at all levels of automation. It is approved for international communication standard used worldwide in the building management systems (BMS). Originally, it is designed specifically for HVAC systems, providing heating, ventilation and air-conditioning of buildings. Subsequently successfully is applied in lighting automation, access control, fire alarm, monitoring and control of elevators and others.It is used to transfer data from network devices, such as pumps, valves, boilers, chillers, fans, sensors, ect.

Using BACnet protocol, building manager can control network devices from distance, including to start / stop them, change operating modes and settings. Devices plugged into the network offer multiple services such as data transfer, management of schedules, reporting on current consumption, alarms and more. That allows building manager to continuously optimize its operation.

BACnet standard supports various media data that defines the physical basis of each network. This makes it compatible with both expensive high-speed cabling, as well as less expensive and slow networks.[4]

Open source network platform LonWorks (Local Operating Network) is created by Echelon Corporation in 1988. As BACnet it also provides communication between many devices on a network dedicated especially on management applications with a wide range of operation, high reliability, easy installation and maintenance. Its communication protocol called LonTalk is also adopted as an official standard for network management. In building automation, LonWorks is used very successfully to control ventilation and air conditioning, lighting of streets and highways, smart measuring of electricity, lighting control and control of the sound of large buildings (stadiums, concert halls, etc.), fire alarm systems, access control, home automation, etc.

LonWorks, as each other local operating network, includes smart devices or “nodes”, interconnected by one or more transmission settings and it communicates using a common protocol. Devices that use LonWorks are equipped with a chip on Echelon. Nodes are programmed to send messages to one another according to the changes in different conditions and perform certain actions in response to messages received. Although the function of individual node sometimes is quite simple and the device is relatively simple (light switch,

temperature sensor, motion sensor, etc.), the interaction between nodes allows LON to perform many complex tasks.[5]

On one hand, specialists involved in the management and automation of buildings (facility managers, engineers, architects, etc.) disagree whether it is better to apply BACnet or LonWorks. BACnet protocol is completely independent and does not require devices to be equipped with a chip on one or another manufacturer. However, LonWorks has been widely used. On the other hand, many manufacturers of automation solutions for buildings provide opportunities to its systems to communicate in both protocols in different types of networks.

Another open source communication protocol is EIB (European Installation Bus). It provides automatic monitoring, management and control of devices and processes, such as lighting, blinds, heating, ventilation and air conditioning, load management, signaling, monitoring, alarms and more. The principle of working is that, the sensors such as motion detectors and thermostats are sending impulses through the transmission setting to the actuator. The basic advantage of this system is that, it can control up to 65,536 devices.

In 1999, as a result of building automatic systems improvement, three European standards (EIB, European Home System - EHS and Batibus) merged into one communication protocol KNX. Nowadays, it is the most common protocol worldwide. An important advantage is that more than 200 companies produce products compatible with KNX. That allows integrators and users to combine products from different manufacturers in one project, although in cases one of them does not entirely satisfy their requirements. Another feature of the system based on KNX protocol is that they are decentralized – a central control unit (controller) lacks. Control is distributed evenly among all the actuators and sensors. This provides outstanding safety and reliability of operation, since a loss of any of these items does not interfere with the work of others.[6]

Finally, the buildings and facilities management is determined by IFMA (International Facility Management Association) as a practice of coordinating the physical workplace, people and activities in an organization. Also, the definition includes the principles of business administration, architecture, behavioral and engineering sciences. Thus facility management is practiced by businesses themselves.

Modern industrial and building automation is unnecessary and unpractible without the use of network technologies. Their widespread distribution is determined by several facts. Among them is the introduction of intelligent devices as a qualitatively new stage in the development of automation. Today, measuring technology parameter is more than the transmission of a certain value. Smart transmitter provides not only many opportunities, but also requires significant information exchange. Similar is the situation with intelligent actuators. Programmable controllers (PLCs) have created a basis for the deployment of advanced networking technologies. Like the digital devices they provided flexibility to the system and opened new horizons for many automation solutions. Last but not least, the search for optimal methods of communication, cabling and reliability helps for development of the automation networks.

References

1. Intelligent Buildings – enabling technologies and the benefits of their implementation, CIO magazine, issue 2, 2008.
2. Design and management of the network infrastructure in the intelligent building, CIO magazine, issue 2, 2009.
3. Bedrov, J., Project of building automation system (BMS), Explanatory Memorandum, S., 2009.
4. Thinking Buildings Universe, http://cbs.grundfos.com.
5. Engineering Review Magazine, issue 3, 2008.
6. Building automation systems – decentralized and based on KNX protocol, Inzhener.bg, 2013.

Mobile Financial Services - Important Prerequisite for the Mobile Business

Violeta Kraeva

D. Tsenov Academy of Economics
kraeva@uni-svihtov.bg,

Petya Emilova

D. Tsenov Academy of Economics
petiaem@uni-svishtov.bg

Abstract: Mobile financial services have great potential for the development of mobile business and becoming more widespread and popular. They are the key to success in today's mobile world. The report analyzes the types of mobile financial services and their advantages as innovative banking services. Some trends of their application both globally and in Bulgaria are pointed out. The mobile system for electronic payments SEP Bulgaria and Bulgarian new mobile platform Mobb are presented.

Keywords. e-commerce, research approaches, trends of development

***

With the expansion of mobile technology in public life, mobile financial management gets more and more widespread. It facilitates extremely much the consumers and revolutionizes the way they manage their monthly expenses, bank accounts or commercial operations. Most of the commercial financial institutions invest in integrated mobile and online platforms. So far, 38% of the banks offer the banking through tablet option, and the creation and maintenance of a competitive mobile channel is a priority for 81% of U.S. banks [1].

The report of the European Payments Council (EPC) from the middle of 2010 confirms that the mobile channel is an ideal platform for the development of payment instruments and forecasts, while the analyst firm Berg Insight suggests that the global number of users of mobile banking will grow to 913 million in 2014, while in 2008 they were only 20 million [2].

According to the results from the MasterIndex Bulgaria survey, conducted regularly by MasterCard, the mobile payments have great potential in Bulgaria - 61% of users aged under 35 and 36% over 35 would use them. 53% of them would pay household bills by mobile phone [3].

The main prerequisites for this are the widespread use of mobile phones (especially smartphones) and mobile Internet. More often, applications are used for payment of bills through mobile phones, receiving SMS for completed transactions in bank accounts and cards, and checking email as well.

Mobile financial services can be divided into five sectors:

- Banking operations;
- Brokerage services (brokerage);
- Investment activities;
- Pay bills online;

- Other services, such as insurance, financial consulting, operations with credit cards and even providing loans and credits.

Some European banks contract providers of mobile radio networks for creating **mobile banks** dealing exclusively with mobile transactions and banking operations.

Entertainment is one of the fastest growing trends in mobile services. In Japan, which holds the leading position in the high-tech, 40% of the traffic of mobile access is occupied by entertainment services, which include cards, jokes, comics, horoscopes and different types of network games.

With the increasing popularity of mobile phones and their massive and widespread use there is an opportunity to introduce a new type of service - mobile payments through short text messages (SMS), sent by mobile phone. As a result, the mobile telephone is converted into individual personal terminal to access electronic services and transactions.

Mobile payments are the result of the convergence of information and telecommunication technologies with banking operations. They can be regarded as a development of the electronic payment instruments and their modern alternative. The first attempts for mobile banking were made in the U.S. in 2000, but merely in 2007 one of the most popular banks - Bank of America, introduces it on a mass scale for about 500,000 customers [4]. This mobile solution is supplemented by a special program of Microsoft MapPoint, which is designed to show users the location of the nearest bank branch or ATM. Mobile payments are widespread also in Canada, Japan, Singapore and other economically developed countries.

In order to impose mobile payments on a great scale, some **conditions** are necessary, as appropriate infrastructure and technology for high-quality wireless data transfer in a short distance (NFC - Near Field Communication) [1]. It is also important the business to realize the significance, speed and convenience of mobile technology in order to implement them. Willingness of people to use mobile payments also affects undoubtedly the spread of these innovative methods of payment. Mostly they need to be convinced in their security and safety.

In the practice, there are **three** ways for mobile payments:

- **mobile commerce (m-commerce)**, which uses mobile browser and online portfolio. The most popular mobile browsers are Safari, Android, Opera Mini, Google Chrome and others;
- **mobile payments (m-payments)**, using mobile applications, eg. of PayPal;
- **mobile wallet (m-wallet)**, which replaces the full payment in cash, for example. Google Wallet

The mobile phone (smartphone) can be also used for other innovative banking service - **Mobile Banking**. It allows via your mobile phone to check balances on bank accounts, to order transfers and payments, to open new accounts and more. So a difference should be made between mobile banking and mobile payments. In fact, the trend is towards a gradual consolidation of mobile banking and mobile payments into mobile wallet.

The main **advantages** of mobile payments could be summarized as follows:

- **flexibility and rationality** because they combine all credit and debit cards in one place;
- **interactivity** by providing the opportunity for personal financial management and offering suggestions for optimal monitoring and controlling of personal finance;

[1] Devices with embedded NFC can read and transmit data to other similar device near 10 cm, as the scope of the action is.

- **speed and security**, because they save time in the dynamic lifestyle and are comfortable and practical not only for consumers but also for businesses and banks;
- **online control** of personal finances anytime and anywhere.

For the first time on a mass scale the innovative mobile payments were used during the Olympic Games in London in 2012 by the payment system Visa. The majority of them were for purchases in shops, services and entertainment, as well as tickets for competitions, marketed at the last moment and passes for theater.

Many suppliers of commercial services offer readers for mobile credit cards that are attached directly to the smartphone. They function like a POS credit card terminal in the store, but are portable and require appropriate application. Worldwide, one of the most popular mobile processors for credit cards is Square.

Generally mobile payments are still in the initial phase of distribution, but growth rates are too high. Just for one year from 2011 to 2012, mobile payments increased 4 times [5]. At the same time PayPal mobile payments were made for 14 million dollars. The forecast is that till 2017 the total value of global transactions conducted through mobile devices will reach about 1.5 trillion dollars, which compared to $ 120 billion in 2012 it is over tenfold more.

The number of consumers using mobile payments will also increase in a very rapid pace and is expected by 2017 to surpass 500 million. That's more than five-fold increase compared to the end of 2012. This undoubtedly proves that mobile payments are the most modern and dynamic electronic financial service. There are even bolder forecasts. According to the Pew Research report "The Future of Money" most likely the money will cease to exist in 2020 when people will fully utilize the use of smart mobile devices for shopping and will eliminate the need for banknotes and credit cards.

Mobile features, that financial institutions currently offer are the traditional banking services to check account balances, review banking transactions, perform money transfers and payments, manage personal finances, purchases by giving credit at the time of sale, etc.

Bulgaria already brings mobile payments as a new distribution channel for e-services via mobile phone. The national e-infrastructure for electronic payments includes free with no financial or technical constraints leading banks, mobile operators, retail and utility companies. For the development of an integrated payment system based on mobile communications "SEP Bulgaria" (Electronic payment system) company was established, which acts as the third operator of the payment system in the country after "Bankservice" and BORIKA. The system provides equal access for all users, ensuring full compatibility and high security. It shall adopt a common standard that turns mobile payments into a universal medium of exchange between consumers and traders, regardless of the mobile operator and bank serviced. This feature is a prerequisite for making mobile payments a viable alternative to cash payments.

In SEP Bulgaria joined the mobile operators Globul and Vivacom and Allianz Bank Bulgaria and EFG Eurobank (Postbank) are involved as banking institutions. Forthcoming accession of DSK Bank as a participant in this market, will make mobile payments even more popular. From the side of the businesses participants are the retailers Technopolis, Piccadilly, Kenar, GEOX, KFC and others. It is envisaged the adoption of mobile payments in restaurants and hotels with the option of splitting the bill, for which purpose, is needed the integration with cash register software used by them. This will account the actual turnover of the restaurant and the business will go in the "daylight".

The main benefit for users of the new system is the ability to transform the mobile phone into a universal tool for personal confirmation of all kinds of electronic transactions and payments in a mobile environment. The technology is based on the use of personal electronic

portfolio that provides online connection to the banking institutions to authorize electronic payments. Through its portfolio each user is able to manage up to five bank accounts in different banks, and makes secure access to various e-services. The user's connection to the electronic portfolio of mobile phone is done easily and conveniently by selecting from a menu of commands and services through confirmation with PIN.

A major **benefit for businesses and public organizations** is the ability to access new distribution channel of services and payments, without incurring any risks to its development and maintenance. The system provides the infrastructure needed for online connection between consumers and merchants and their banks. The new channel is expected to gain widespread popularity among consumers due to the massive mobile penetration and a set of comparative advantages.

Based on open business model between banks, mobile operators and retailers, equal access to services and resources of the payment system is ensured. Furthermore, for the first time the customers operations are carried out from their mobile phones authorized and implemented in real time, while at BORIKA for example, they are authorized online, but the link with the bank server is done twice a day.

In October 2013 was the beginning of the new Bulgarian mobile platform of Borika-Bankservice called **Mobb**. It provides passive informational and active payment services to customers using smartphones and various bank cards without requiring data entry for the cardholder and his card. This is a unique, reliable and secure mobile service that is very convenient and will imposed itself as a necessary tool for mobile payments in the everyday life of customers.

***

In conclusion, we could emphasize that mobile business applications are no longer a luxury, but are becoming in a priority and a necessity for modern enterprises from various industries and financial institutions. As some experts say, thanks to wireless communications, access to most of the corporate network is in the pocket of its employees. So the mobility increases business productivity through greater flexibility in routine business processes regardless of their place and time. It can be seen as the next step in the natural evolution of the business and its role will increase more in the coming years. In response to this trend, banks need to develop a clear mobile strategy and new basic mobile business model in order to be competitive, flexible and leading in today's mobile business.

References

1. Какво да очакваме от банките на бъдещето. CIO, март 2013, с. 63.
2. **Иванов, Т.** За електронните плащания и хората. Computerworld, бр. 1-2, 15.01.2011, с. 8.
3. http://www.etimes.bg/articles/1011/Mobilnite_plashtaniq_s_golqm_potencial_u_nas.html (26.01.2011).
4. Развитието на електронните плащания зависи от състоянието на банковата система. http://review.sagabg.net/item_8042.html (21.11.2013).
5. Защо мобилните плащания са обречени на растеж. http://infozone.bg/node/31 (11.12.2013).
6. **Steffano Korper and Juanita Ellis**. The E-Commerce Book (Building the E-Empire). Academic Press, 2001, ISBN: 0-12-421161-5
7. **Urbaczewski, A., L. M. Jessup, B. Wheeler**, "Electronic Commerce Research: A Taxonomy and Synthesis," Journal of Organizational Computing and Electronic Commerce, Vol. 12, No. 4 (2002), pp. 263–305.

Application of Time Series Analysis of Competing for a Limited Resource Economic Systems

Miroslava J. Ivanova[1], Ivan P. Jordanov[2], Nikolay K. Vitanov[3]

[1,2] UNWE, 1700 Sofia, Bulgaria
[1]mi_ka@abv.bg, [2] i_jordanov@email.bg

[3] Institute of Mechanics - BAS, Acad. G. Bonchev Str., Bl. 4, 1113 Sofia, Bulgaria

Abstract. Most of the systems in the world around us contain elements that influence one another through competition and cooperation. Examples of such systems are the different economic organizations (Dimitrova and Vitanov 2005). The article explore time series derived from a model of the dynamics of competing for a limited resource economics entities These time series are stationary and their length can be arbitrarily chosen so that to apply different methods for analysis of stationary time series. We choose these time series so that the system is chaotic attractor. Examine the time series in summary Volterra systems and found intermittent behavior research model for periodical, chaotic and poorly chaotic behavior of the system. We confirmed the conclusion of the theory of Zauer, York and Kasdagli for the dimensionality of the phase space.

Keywords stationary time series, nonlinear dynamics, Dimitrova – Vitanov model, Volterra equations, attractors

1. Introduction

Development of humanity leads to a continuous increase in the number of people which is a factor in increased consumption of natural resources. We see more and more disasters further by inappropriate human activities. Therefore in science naturally the task of understanding the dynamics of the processes in the societies of many individuals, including processes in sociology [2,3,4] and processes in the economy [1].

Much of the complex economic and social systems contain elements that influence each other through competition or cooperation [11, 12, 13]. With the increasing complexity of this interaction, competing elements begin to adapt to environmental changes. In recent years, methods of nonlinear dynamics, chaos theory and time series analysis began to be used increasingly in the study of such adaptive systems. The main features of such complex systems can be understood on the basis of a relatively small number of non-linear differential equations, which in the above assay methods are particularly suitable. Here we examine the time series related to the dynamics of competing for a limited resource entity.

2. Methodology to the study of time series

The study of time series can be performed with statistical quantities such as average order values of $\boldsymbol{N}$, the standard deviation or other linear features. This linear characteristic

is the power spectrum $P(w)$ which is obtained by Fourier transformation of the magnitude $u(t)$:

$$F(\omega) = \int_{-\infty}^{\infty} dtu(t)e^{i\omega t} \tag{1}$$

The presence of a dominant frequency and its harmonics is just a sign of cyclic of the test system and the existence of several dominant frequencies associated with multi-periodic motion.

Note that chaotic attractors are characterized by several peaks in the spectrum of power, accompanied by a band of power concentrated in the low frequency region of the spectrum ω. Namely the presence of such dominants frequencies are in the base of the methods for predicting the of the time series. Another linear feature is histogram $\rho^*(x)$ when the values of the time series are at regular time intervals, then the corresponding value of $\rho^*(x)$ are the proportion of time spent in the range of about x.

Particularly important information about the evolution of the system gives us the autocorrelation function:

$$A(l) = \int_{-\infty}^{\infty} dtu(t+l)u(t) \tag{2}$$

The above function is periodic for periodic signal. For deterministic chaotic system, however, this function falls exponentially with increasing pitch.
Thus is time-delay reconstruction of the dynamics of the systems [5,6]. One of the most popular to the reconstruction of phase space is [16,17]. From the sequence data $\{x_n\}, n = 1,2,3,...$, recorded at the time t_n construct vectors:

$$\vec{\mathrm{x}}_{\mathrm{k}} = \left(x_{k-(m-1)r}, x_{k-(m-2)r}, ..., x_{k-r}, x_k\right) \tag{3}$$

where m is the dimension of the resulting vector space, and τ is the step between two sweep. Has been shown (Sauer 1991) that most real systems properties embedding occurs when the dimension of the reconstructed phase space is two times greater than the box counting dimension of attractors. When we examine stationary time series is convenient step τ and the dimension m to select using the False nearest neighbors [16,17].

3. Obtain time series from the model Dimitrova - Vitanov

Time series are derived on the basis of a model about the dynamics of three populations competing for limited resources. The basis of this model is the system of equations used by Abramson [7] in modeling the extinction of populations and their replacement by other populations:

$$\frac{dN_i}{dt} = N_i \left(f_i + \sum_{j=1}^{n} b_{ij} N_j \right), i = 1, 2, \ldots, n \tag{4}$$

In these model equations f_i is so called a fitness member describing the ability of a population to multiply when left to itself, without contact with other populations. The coefficient b_{ij} describes the interaction of i-th and j-th populations. In classical models the fitness and the interaction coefficients imply constant over time. This is a special case of the proposed by Dimitrova and Vitanov model (Dimitrova and Vitanov 2005), in which the fitness and the coefficients of interaction are assumed dependent on the number of individuals in the population. The first few members in these relationships look like this:

$$f_i = f_i^0 + \sum_{k=1}^{n} f_{ik} N_k + \sum_{k=1}^{n}\sum_{l=1}^{n} f_{ikl} N_k N_l + \ldots \quad (5)$$

$$b_{ij} = b_{ij}^0 + \sum_{k=1}^{n} b_{ijk} N_k + \sum_{k=1}^{n}\sum_{l=1}^{n} b_{ijkl} N_k N_l + \ldots \quad (6)$$

Let break the above two developments since their second members, so get convenient for analysis system with variable fitness and coefficients of interaction between populations. We substitute the obtained result in the system (4) and reach out to a common model system of ordinary differential equations:

$$\frac{dN_i}{dt} = N_i \left[f_i^0 + \sum_{j=1}^{n} \left(f_{ij} + b_{ij}^0 \right) N_j + \sum_{j=1}^{n}\sum_{k=1}^{n} b_{ijk} N_j N_k \right] \quad (7)$$

Our goal in this article is to investigate chaotic time series generated by the system (7). This is done by the system are selected so that the attractor of the system is been chaotic through the choice of the coefficients b_{ij}^0:

$$b_{ij}^0 = \begin{pmatrix} -k_1 & -k_1 & -k_2 \\ k_1 & k_2 & -k_2 \\ -k_3 & -k_2 & -k_2 \end{pmatrix} \quad (8)$$

Additionally, we minimize the remaining parameters of the system like we require:

$$b_{ijk} = b, \quad i,j,k = 1,2,3, \quad f_{ij} = f, \quad i,j = 1,2,3 \quad (9)$$

$$f_1^0 = 2k_1 + k_2, \quad f_2^0 = -k_1, \quad f_3^0 = k_3 + 2k_2 \quad (10)$$

Thus keeping down the number of independent parameters of the system to the 5 - k_1, k_2, k_3, b and f and the system of equations takes the form:

$$\frac{dN_1}{dt} = N_1 \left[2k_1 + k_2 + \left(f - k_1 \right) N_1 + \left(f - k_1 \right) N_2 + \left(f - k_2 \right) N_3 + T \right] \quad (11)$$

$$\frac{dN_2}{dt} = N_2 \left[-k_1 + \left(f + k_1 \right) N_1 + \left(f + k_2 \right) N_2 + \left(f - k_2 \right) N_3 + T \right] \quad (12)$$

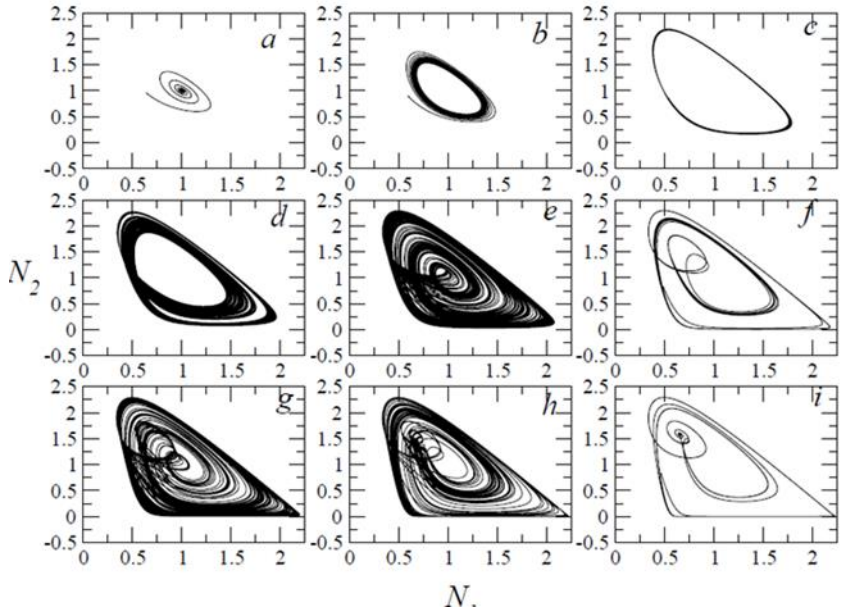
$$\frac{dN_3}{dt} = N_3 \left[k_3 + 2k_2 + \left(f - k_3 \right) N_1 + \left(f - k_2 \right) N_2 + \left(f - k_2 \right) N_3 + T \right] \quad (13)$$

where the member T has the form:

$$T = b\left(N_1^2 + 2N_1N_2 + 2N_1N_3 + N_2^2 + 2N_2N_3 + N_3^2 \right) \quad (14)$$

The chaotic behavior of the studied system is Shilinkovsik type (originators of the study Shilinkovski chaos in systems of competing populations are scientists from the group around Alain Arneodo [8,9,10].

For the realization of this relatively rare type of chaos requires two fixed points the P_1 and P_2 of the studied system, as one of them should lose their stability through local bifurcation. In this case, the local bifurcation we need is bifurcation of Hopf (investigated also by Poincaré and Andronov) and stationary point which becomes unstable through this bifurcation is the point P_2. By using the theorem of the local variety, the occurrence of bifurcation of Hopf and the initial stage of its development may to explore analytically. We'll leave those huge volume estimates of colleagues mathematicians.



4. Analysis of the obtained time series

Surveyed time series are obtained by numerically solving the system of equations (11-14) to $k_1 = 0.1, k_2 = 0.5, b = 10^{-3}$ and $f = 10^{-6}$. Values for k_3 are respectively $k_3 = 1.1$, $k_3 = 1.43$, $k_3 = 1.54$. The behavior of these time series is clearly seen from the figure above. Note that we have chosen the time series so that one of them has a periodic behavior, the second to have erratic behavior, and the third to have a lot less chaotic behavior. To calculate each of the histograms divide the interval change N_3 of 100 equal parts. Yet here we see that the interval of climate variables studied is different for different values of the parameter. Unlike other case of examination of the time series of integrated systems of Volterra [13, 14, 21, 22, 23], instead of several peaks in our histograms have only one peak at small values of N_3. This shows us that the system spends most of its time in

small N_3 and only briefly the value of N_3 is increasing. Thus, behavior is widely spread further and was called intermittent. In this intermittent behavior is different for different values of k_3 as illustrated by figure.

Now compute the autocorrelation function $A(n)$, where n is the number of steps between the two values for which the correlation was calculated. We find that when $k_3 = 1.1$ autocorrelation function is periodic. For the chaotic system, however, the autocorrelation function decays very rapidly (almost exponentially) to 0. By increasing the periodic components increases range correlations.

Finally, we calculate the number of false neighbors for the three investigated time series. Where the percentage of false neighbors becomes low is considered that the embedding dimension is sufficient for reconstruction of our phase space.

For our case, we get that the minimum adequate dimension at the reconstructed phase space is $m = 3$, $k_3 = 1.1$ and $m = 7$, $k_3 = 1.43;1.54$. Thus confirming the conclusion at the theory at Sauer, York and Kasdagli for the minimum dimension at the reconstructed phase space is a 2 times +1 larger than the fractal dimension at attractor.

In this case, the dimension at attractor for $k_1 = 1.1$ is 1, i.e. attractor is a cycle, while the remaining two attractors received three dimensions, and i.e. they are chaotic.

5. Conclusion

In this article, using different methods for the analysis of time series, we received additional features of the dynamics in phase space for a model systems of ordinary differentials equations associated with the dynamics of interacting (competing or co-operative) for limited resources systems. These equations with appropriate choice of parameters to be reduced to a well know models from biology, physics, economics and sociology. Although the model study changing the value of only one parameter , we got the full range of qualitatively different attractors in three-dimensional phase space (fixed points, border cycle and chaotic attractor), while chaos is called Shilnikovski, obtained by grip of the cycle (bifurcation of Hopf) of connection associated with losing his truck stability focus. So the resulting time series are stationary, which allowed us to apply the full range of methods of classical linear and nonlinear time series analysis. Of course, we had analytical model differential equations of lines we received with sufficient length, containing information about all the characteristic frequencies of the system.

So we found intermittent behavior generalized Volterra systems with periodic, chaotic and slightly chaotic system behavior. We confirmed the theory at Sauer York and Kasdagli for a dimension of the phase space.

Acknowledgment

We thank to the Nonlinear time series analysis group from the Max-Planck Institute for the Physics of Complex Systems in Dresden, Germany, for the possibility of using of the TISEAN software package.

References

1. Балабанов Илия, Витанов Николай. К., Икономическият метроном, Икономическа мисъл, ЛВ, 6.
2. Витанов Николай К., З. Димитрова, С. Панчев. Популационна динамика и национална сигурност. 2005: Академично издателство "Марин Дринов" на БАН, 2005.
3. Витанов Николай К., З. Димитрова, С. Панчев. Социална динамика без формули. София: Академично издателство "Марин Дринов" на БАН, 2008.
4. Димитрова З. И., Н. К. Витанов, Популационната динамика и неочакваните и приложения. Наука, 12, 27.
5. Abarbanel H. D. I. Analysis of Haotic Time Series. Springer, New York, 1996.
6. Abarbanel H. D. I., R. Brown, J. J. Sidorowich, L. Sh. Tsirming. The analysys of observed chaotic data in physical systems. Rev. Mod. Phys. 65, 1331-1392, 1993.
7. Abramson G. Ecological models of extinction. Phys. Rev. E 55, 785
8. Arneodo A., P. Coullet, J. Peyraud, C. Tresser. Strange attractors in Volterra equations for soecies competition. J. Math. Biology 14, 153. 1982.
9. Arneodo A., P. Coullet, C. Tresser. Occurrence of strange attractors in three-dimensional Volterra equations. Physics Letters A 259-263, 1982.
10. Arneodo A., P. Coullet, C. Tresser. Oscillators with chaotic behavior: An illustration ofa a theorem by Shilnikov. J. of Statistical Physics 27, 171-182, 1982.
11. Dimitrova, Z. I., N. K. Vitanov, Influence of adaptation on the nonlinear dynamics of a system of competing populations. Phys. Lett. A 272, 368-380, 2000.
12. Dimitrova, Z. I., N. K. Vitanov, Dynamical Consequences of Adaptation of Growth Rates in a System of Three Competing Populations. J. Phys. A: Math Gen., 34, pp. 7459-7473, 2001.
13. Dimitrova Z. I., N. K. Vitanov, Adaptation and its impact on the dynamics of a system of yhree competing population. Physica A, 300, 91-115, 2001.
14. Dimitrova, Z. I., N. K. Vitanov, Chaotic pairwise competition, Theor. Population Biol. 66, pp. 1–12, 2004.
16. Hegger R., H. Kantz, Improved false nearest neigbours method to detected determinism in time series data. Phys. Rev. A 60, 4970, 1999.
17. Hegger R., H. Kantz, T. Schreiber. Practical implementation of nonlinear time series methods: The TISEAN packade. CHAOS 9, 413, 1999.
18. Jordanov I. P. On the nonlinear waves in (2+1)-dimensional population systems. Comp. rend. Acad. Sci. Bulg. 61, 307-314, 2008.
19. Jordanov I. P. Nonlinear waves caused by diffusion of population members. Comp. rend. Acad. Sci. Bulg. 62, 33-40, 2009.
20. Vitanov, N. K., On Modified Method of Simplest Equation for Obtaining Exact and Approximate Solutions of Nonlinear PDEs: The Role of the Simplest Equation. Commun Nonlinear Sci Numer Simulat, 16, pp. 4215–4231, 2011.
21. Vitanov, N K., Z. I. Dimitrova, Application of the Method of Simplest Equation for Obtaining Exact travelling-wave Solutions for Two Classes of Model PDEs from Ecology and Population Dynamics. Commun Nonlinear Sci Numer Simulat, 15, pp. 2836–2845, 2010.
22. Vitanov, N. K., I. P. Jordanov, Z. I. Dimitrova, On nonlinear dynamics of interacting populations: Coupled kink waves in a system of two populations, Commun.Nonlinear Sci. Numer. Simul. 14, pp. 2379–2388, 2009.
23. Vitanov, N. K., E. Yankulova, Multifractal analysis of the long-range correlations in the cardiac dynamics of Drosophila Melanogaster, Chaos, Solitons Fractals 28, pp. 768–775, 2006.

Using SPSS for Process Quality Control – A Critical Review

Alexander Naidenov,
Statistics and Econometrics Department, UNWE, Sofia, Bulgaria,
anaidenov@gmail.com

Abstract. The pursuit of high quality production is one of the main topics of our modern world society. The great variety and complexity of the production characteristics lead to the need for use of the specific software products in order to control for the production process quality. There are many software solutions to this issue but one of the most famous is SPSS. Although the latter is a wide range purposed software tool it also provides great opportunities for a basic quality control analysis. The paper is a critical review of SPSS quality control functions and features.

Keywords: SPSS, quality, control, process, tools

1. Introduction

Our contemporary society is dedicated to the mass production and consumption of incredible wide variety of products. Some manufacturers prefer to produce low cost products with low quality but other pursuit the excellence in all directions.

When we talk about quality products we understand different things. The consumer imagines a quality good or service as such that would satisfy his/hers physical, emotional and mental needs. But if we ask the manufactures they will replay: 'Our quality products (services) are such because they meet the required standards'. When these two viewpoints collide the concept of production process quality is born.

The manufacturers who follow their production standards and strive to satisfy consumer needs are obliged to observe a great number of rules concerning the product or service characteristics and features. In order to achieve a certain level of 'perfection' in these quality characteristics (e.g. weight of a biscuit, length of a knife blade, smoothly working tablet PC operating system, etc.) we need to probe continuously the status of these. This is due to the fact that principle 'everything changes' holds for the production characteristics too. The variation of the latter is caused by different kind of sources e.g. tired worker makes more mistakes, worn machine produces poor output, bad raw materials lead to bad products, etc. So if we can reduce this variation in the characteristics we will achieve easier the standard requirements and finally we will have production with certain level of quality.

The first step to manage the variation of production characteristics is to survey and observe their changes in time. If we are able to measure and control the levels of these characteristics we will be able to provide assurance of good production quality.

There are many statistical and non-statistical tools for production process quality control but one of the most popular ones, even now - ninety years after their invention, are so called Shewhart control charts. Generally speaking they are a special graphical visualization of certain product characteristic temporal dynamics. These charts provide the possibilities for observation of characteristics' changes in the time and interference in the production process, if there is a warning for product quality worsening.

Unfortunately one product or service doesn't have only one dimension of quality. Even the one-piece type products, such as the kitchen knife, have more than one characteristic – blade length, hardness, sharpness, handle comfort and etc. In order to control

for all the quality features of a given product, we usually use modern technologies such as computers, sensors, lasers and so on. Nowadays every product, subject to quality control, is an object of detailed measurements which are stored as a raw data in big databases. The latter are then imported in specialized software products to deepen the analysis of the process dynamics and variation. Many of today's statistical packages, which are multi-purposed ones, include a set of statistical methods for quality control of production process data. This paper considers the most famous and widely used software for broad-range analysis - Statistical Package for Social Science (or SPSS) and its usage for basic quality control analyses. Next sections are dedicated to a short description of the software capabilities in the quality control field and also a critical review of its features.

2. SPSS quality control features and functions

Nowadays many scientists and researchers in Bulgaria count on SPSS for their analysis and problem solving and that is why it is a logical reason to test the capabilities of this software for quality control problem solvation too.

Even though initially SPSS is social science oriented software, over the course of time it became a Swiss-army-knife type tool. Its capabilities include manipulating raw data from different sources and many formats, basic statistical descriptive analysis (including graphical presentation), advanced techniques such as non-parametric tests, regression, correlation, factor analyses, multiple imputation, distribution simulation, sampling, survival analysis and many more. Since version 14th (at the time of writing this paper 22nd version of SPSS is the latest available and this paper is based on it) SPSS team implemented quality control functionality and became an easy tool for basic quality control analysis. Let us present its quality control features at the following diagram:

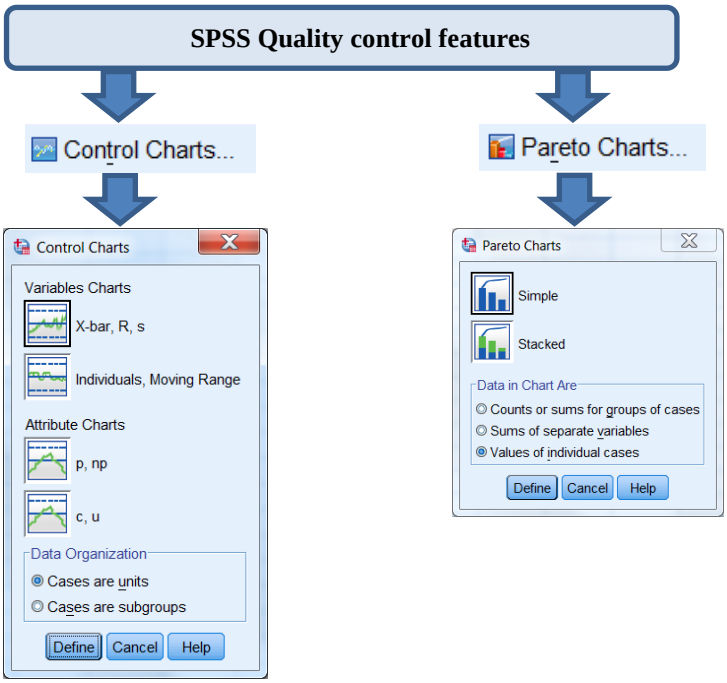


Fig.1. Main SPSS quality control features.

As it can be seen in the Fig.1, SPSS has two main tools for quality control: *control charts* and *Pareto charts*.

Control charts include such as these for variables (quantitative characteristics measures at interval or ratio scale e.g. temperature in Celsius, weight in grams, length in cm, etc.) and for attributes (qualitative characteristics measured at nominal or ordinal scales e.g. defective and non-defective items).

Measurements of a given product characteristic is based on a random selection (a.k.a. sample) of one individual item (unit) or a subgroup (sample) of items (units) from the all produced items called a lot. The reasons for the unit sampling are rooted in the great number of produced items (usually in hundreds of thousands, even millions) and the lack of resources (time, personnel, space, etc.) for the measurement of all characteristics for given units. Also it is imposed by the "unreal" variation in the individual unit characteristics (individual values variation is greater than sample means variation). Using the measurements of the sampled units and based on the statistical sampling theory [3], we can be sure at a certain level that almost all items (99,7%) will fall in ±3 standard errors (standard deviation divided by squared number of sampled items) from the *target value* (the mean of all measurements). This gives us the possibility by the use of *control charts* (which implement the sampling theory) to test if there are measurements outside these limits (also called *control limits*) and if there are such to act in order to avoid manufacturing of out of specification (the specification limits are defined by the quality standards considering the natural variation (not caused by and special events) in the item characteristics and are described in details in the product documentation) items (e.g. if the knife blade became too short of too long). When we encounter such out of control limit measurements it is a warning that may be there are *special (non-natural) causes* influencing the production process that may lead to a low quality production [4].

All these assumptions are considered in the control chart functionality of the SPSS software. First let us describe the software features concerning the *variable charts*. There are two types of them: *charts for subgroup (sample) of units* and *charts for individuals (*in *Data organization* option you define the structure of the file and not the data collection type*)*.

For subgroups *For individuals*

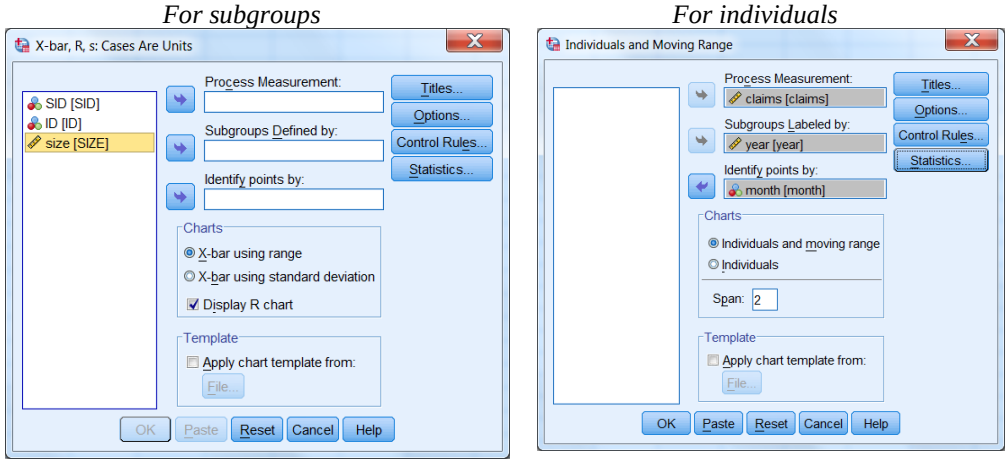


Fig.2. Screenshots of the software windows for variable control charts.

When data are available for a sample of units we use so called *X-bar* (mean), *R* (range) and *s* (standard deviation from a sample) charts. By the combined use of these we can simultaneously account for the change in the *accuracy* (sample means) and for the *precision* (sample range and standard deviation) of a given the production process. In Fig.2 (the left part) we can see the required fields for control chart building. These are: *Process Measurement* (defines the variable that contains the measurements of interest), *Subgroups Defined by (*defines the variable which account for the identification of each sample) and *Identify points by (*defines the variable which identify each and every measurement in order to expose for which of them we should worry about). At this step we can choose how we will measure precision – by R or by s charts. Also there are additional options available for each control chart given as buttons: *Titles* (user defined chart titles), *Options* (defines the number of sigmas (standard error boundaries) to use in testing and the size of each sample (subgroup), *Control Rules* (specific rules used to perform quality control using control charts features [6]) and *Statistics* (control chart capability analysis features using indexes [5]).

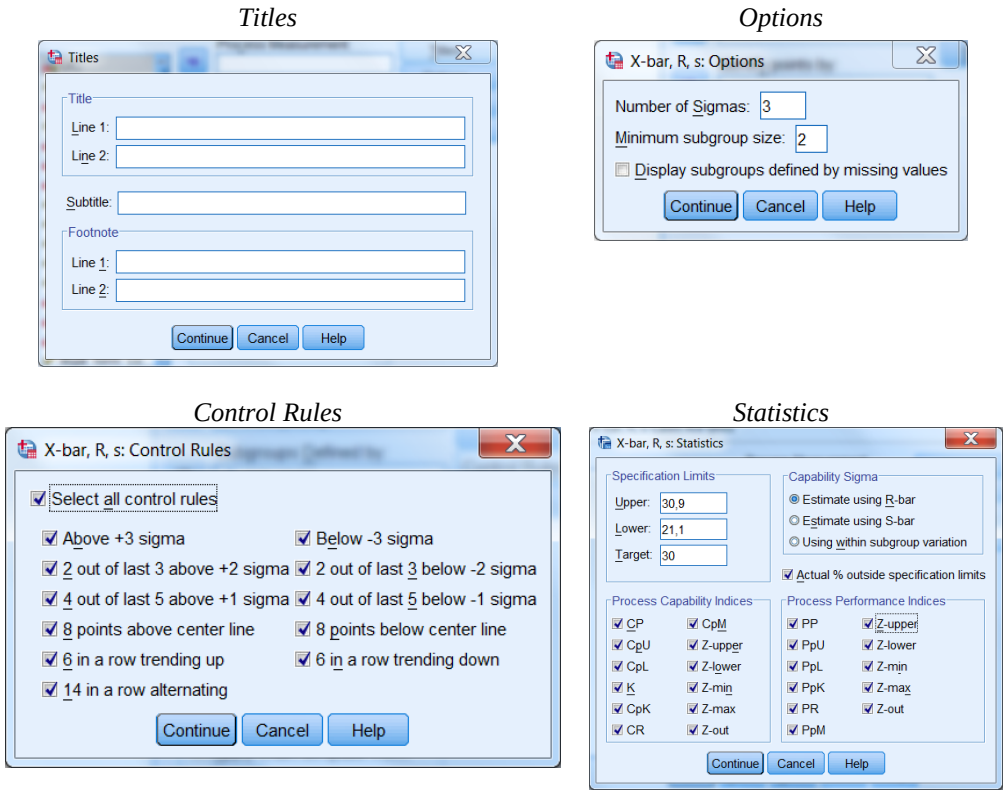
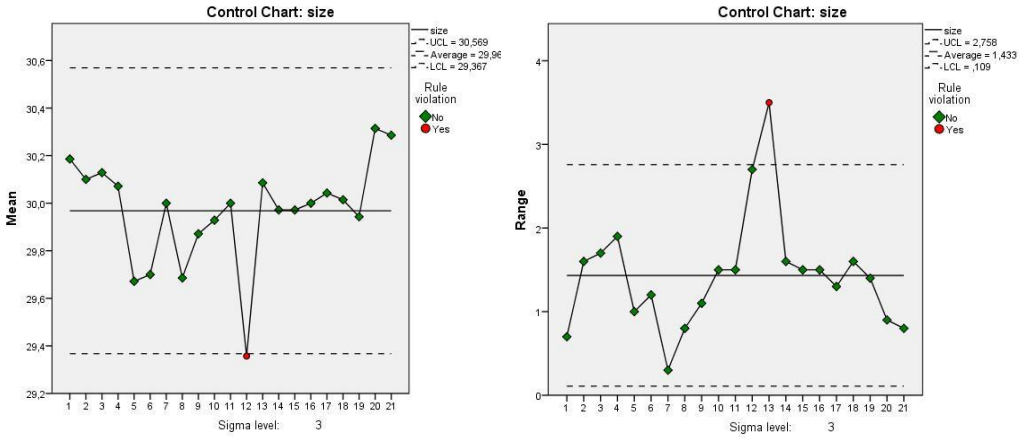


Fig.3. Screenshots of software windows for variable charts options.

Using example data from the production process of terracotta tiles (21 random samples of 7 units) we attain the following *control charts*:

X-bar control chart *R control chart*



Control chart rule violations

Rule Violations for X-bar

ID	Violations for Points
7	Less than -3 sigma

1 points violate control rules.

Rule Violations for Range

ID	Violations for Points
7	Greater than +3 sigma
7	2 points out of the last 3 above +2 sigma

1 points violate control rules.

Fig.4. Screenshots of SPSS output results for variable control charts (an example).

As it can be seen in Fig.4, control charts are presented visually in order to find out if there are any warnings or rules violations. For every chart also we can see the average of sample means (average is given with solid line and means with diamonds), upper and lower control limits (given with dotted lines). Automatically the possible rule violations (presented as circled points) are described in tables too with additional information for the ID of the problematic measurement(s).

An extra feature is the capability analysis which shows the *production process possibilities* in the scope of manufacturing production in the specification limits:

Process Statistics

	Act. % Outside SL	1,4%
Capability Indices	CP[a]	3,082
	CpL[a]	5,577
	CpU[a]	,586
	K	,810
	CpK[a]	,586
	CR[a]	,324
	CpM[a,b]	3,076
	Z-lower[a]	16,732
	Z-upper[a]	1,758
	Z-min[a]	1,758
	Z-max[a]	16,732
	Est. % Outside SL[a]	3,9%
Performance Indices	PP	3,069
	PpL	5,555
	PpU	,584
	PpK	,584
	PR	,326
	PpM[b]	3,064
	Z-lower	16,666
	Z-upper	1,751
	Z-min	1,751
	Z-max	16,666
	Est. % Outside SL	4,0%

The normal distribution is assumed. LSL = 21,1 and USL = 30,9.

a. The estimated capability sigma is based on the mean of the sample group ranges.

b. The target value is 30.

Fig.5. Screenshots of SPSS output results for capability analysis (process statistics).

The features for the *control charts for individuals* are almost the same but considering the fact that here we do not have samples of a number of units we cannot calculate sample means, ranges and standard deviations. In this case we use so called *moving range* or the difference between two consecutive measurements. The use of those kind of control charts is imposed by the peculiarities of some manufacturing or servicing process that have for example only one unit at random time intervals (e.g. customer claims in a bank).

When there are no possibilities for quantitative measurement of production items (e.g. we have only a description of the item quality features) we usually define the units only as defective (non-conforming) and fit-to-use (conforming) ones i.e. we use so called attribute control charts. If we are interested in the number of defects per item then we call these issues non-conformities. In SPSS we can use *np* (number of non-conforming units from all produced) and *c* (number of non-conformities) chart – if each sample used has an equal size, and *p* (proportion of non-conforming units from all produced) and *u* (number of non-conformities per unit) – for the unequal size samples.

p and np control charts *c and u control charts*

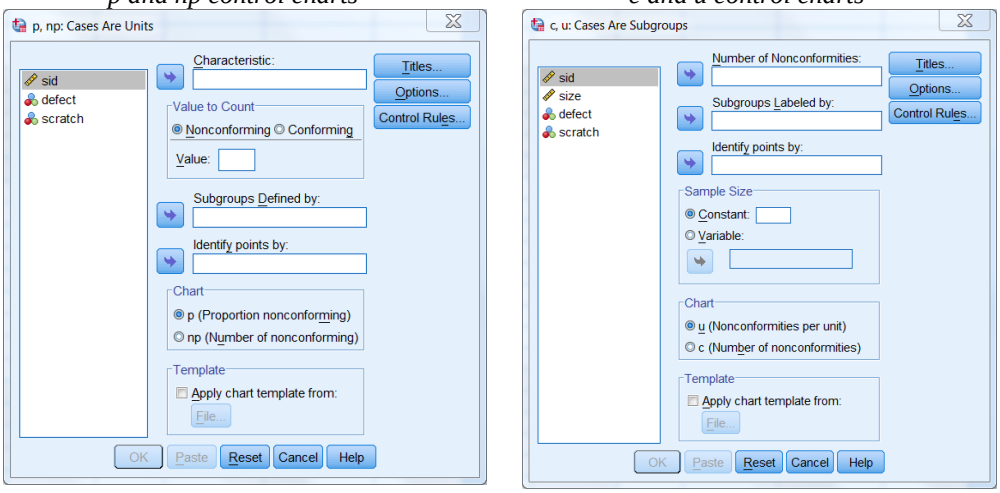
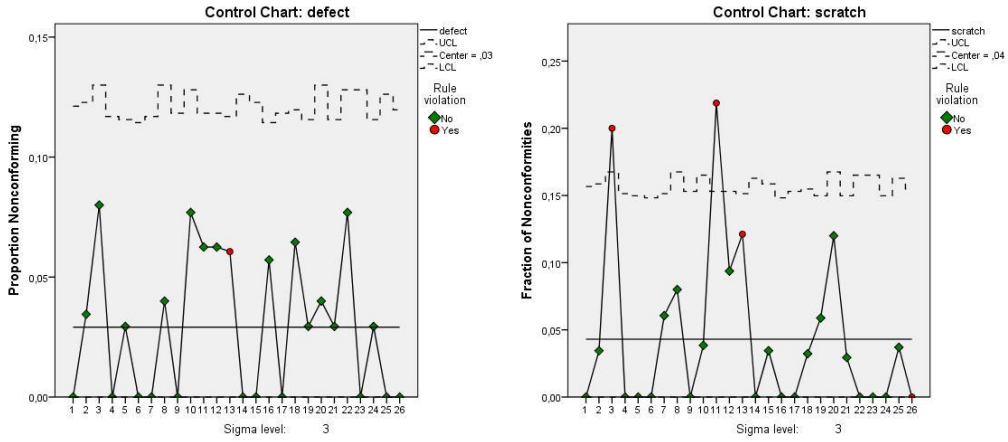


Fig.6. Screenshots of software windows for attributes control charts.

Similarly to the variable control chart we should define: *subgroups* and *point* (units) identifiers but also specific variables for the nonconforming items and the nonconformities. It is important to indicate if we have used equal (constant) or unequal (variable) sample sizes. The options and control rules are exactly the same as those for the variable control charts.

Using sample data for quality control of the color pencil production we obtain the following results for the four types of attribute charts:

p control chart *u control chart*



(cont.)

np control chart *c control chart*

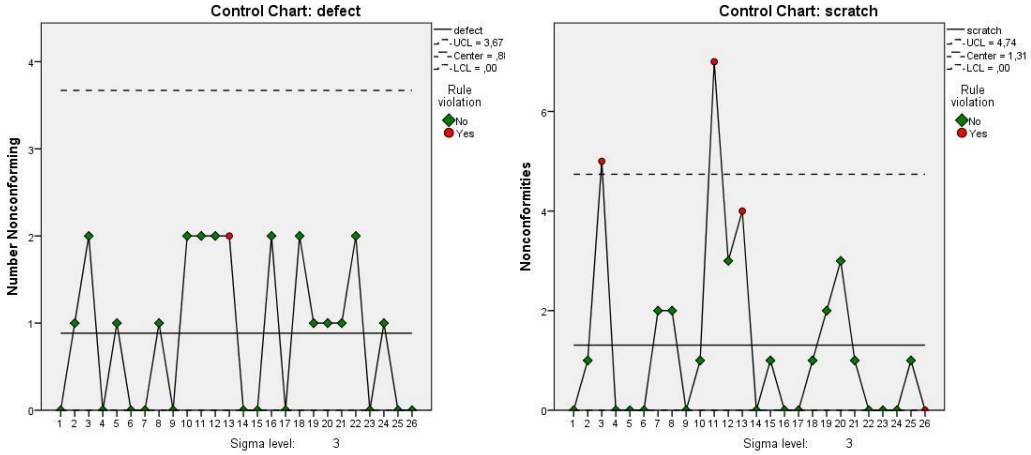


Fig.7. Screenshots of SPSS output results for attribute control charts.

In this case we can also notice the presence of information about the center line (target value), upper and lower control limits (at ±3 standard errors), sample non-conforming items and non-conformities too. In case of need for rules violation analysis an output for these is also available:

For the p control chart

Rule Violations

Case Number	Violations for Points
13	4 points out of the last 5 above +1 sigma

1 points violate control rules.

For the u control chart

Rule Violations

sid	Violations for Points
3	Greater than +3 sigma
11	Greater than +3 sigma
13	2 points out of the last 3 above +2 sigma
26	4 points out of the last 5 below -1 sigma

4 points violate control rules.

For np control chart

Rule Violations

sid	Violations for Points
13	4 points out of the last 5 above +1 sigma

1 points violate control rules.

For c control chart

Rule Violations

sid	Violations for Points
3	Greater than +3 sigma
11	Greater than +3 sigma
13	2 points out of the last 3 above +2 sigma
26	4 points out of the last 5 below -1 sigma

4 points violate control rules.

Fig.8. Screenshots of SPSS output results for attribute control charts (rules violations).

When we deal with attribute data a lack of capability analysis is evident, so in this case it is omitted. However we have additional capabilities for defect items analysis and it is based on so called Pareto charts, named after the Italian economist Vilfredo Pareto (1848-1923) who discovered that 80% of the land in Italy was owned by 20% of the population (see Fig.1 - right part). They are useful if we'd like to find out what are the major reasons for non-conforming units production. For example if execute quality control for a call-center

service, indicating reasons for eventual client claims, we could get the following results by summarizing the collected data:

SPSS Pareto Chart window *Output results*

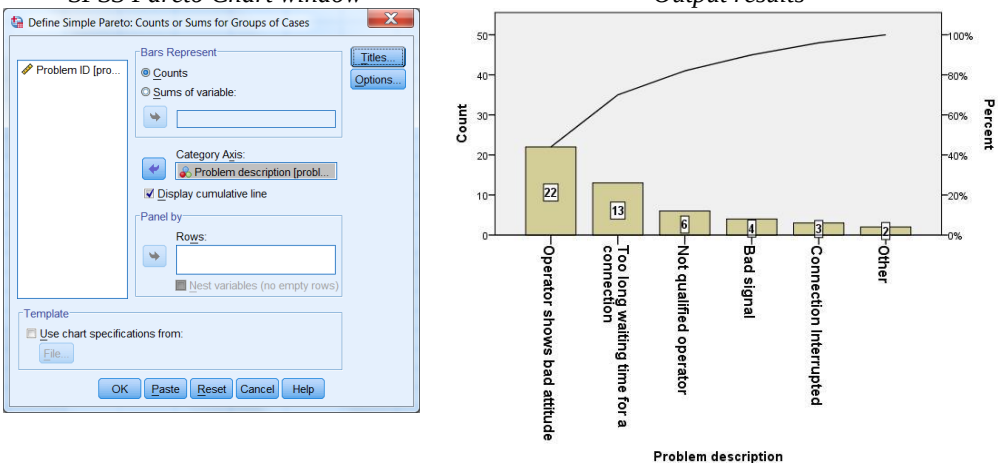


Fig.9. Screenshots of SPSS Pareto charts function and output results.

Depending on the information collected it is possible to use and visualize individual cases or aggregated data. At the figure above (in the right) we could see the Pareto chart for our example that states that the main reasons for call-center issues (see left part of the chart) are the 'Operator bad attitudes', 'Too long waiting time' and 'Not qualified operator'. The chart shows not only the number of claims (counts) but also the cumulative present of different types of issues. We can see that Pareto principle is held that about 80% of the claims are due to the first three reasons.

Although SPSS includes the most frequently used basic tools for quality control it lack some of most powerful techniques implemented in specialized software products such as Statistica and Minitab. Some of these *missing functionalities* are:

- Design of Experiments;
- Cause-and-effect (fishbone) diagram;
- EWMA and CUSUM charts;
- Rare event control charts: G, T;
- Multivariate control charts: T-squared, generalized variance, MEWMA;
- Custom tests for special causes;
- Process capability for non-normal, attribute, batch type of data;
- Six Sigma analysis;
- Acceptance sampling and OC curves;
- Measurement Systems Analysis (Gage Repeatability/Reproducibility Analyses);
- Automatic updating of quality control charts and process capability statistics as new data becomes available (real-time online process analysis and monitoring);
- Sampling techniques for quality control.

The abovementioned functions are not obligatory but they are essential especially when the statisticians or researchers want to do more elaborate analysis of the quality status of a given production process.

3. Conclusion

Even though SPSS is not a 'real' quality control tool it provides a sufficient amount of basic functionalities for the quality issues management. Unlike its qualified and specialized competitors (Minitab and Statistica), which provide sophisticated and elaborated quality control instruments, this software product gives a limited options for quality analyses mainly by the use of control and Pareto charts. However we hope that our favorite universal all-in-one tool for the data processing will continue developing in the scope of its quality control functions and soon we will be able to do more powerful and complex quality statistical analyses.

References

1. IBM Corporation: IBM SPSS Statistics 22 Core System User's Guide (2013).
2. IBM Corporation: IBM SPSS Online Help (2013).
3. Lohr, S.: *Sampling: Design and Analysis, Duxbury Press* (2010)
4. Mitra, A.: *Fundamentals of quality control and improvement*, John Wiley & Sons (2008)
5. Montgomery, D.: *Introduction to Statistical Quality Control, Sixth Edition*, Wiley & Sons Inc. (2009).
6. Oakland, J.*: Statistical Process Control, Sixth Edition*, Butterworth-Heinemann (2008).

Sample Path Analysis of $M^x/G/1/K$ Queue with Thresholds on Queue Length

Mitko Dimitrov
UNWE, 1700 Sofia, Bulgaria
mcdimitrov@abv.bg

Abstract: A single server queue with batch arrivals, finite waiting space, and K thresholds on queue length is considered. Arrivals follow a compound Poisson process. The service times depend on queue length at service-start epoch. The embedded Markov chain at service start epochs and sample path analysis of queue length are employed to investigate the $M^x/G/1/K$ queue with thresholds. The steady state distribution of the number of customers in queue at service-start epochs is obtained. Also the stationary distribution of queue length and remaining service time in the queue is considered.

Keywords: batch arrivals, finite waiting space, stationary distribution of queue length and remaining service time.

1.Introduction

Queuing models with finite capacity and batch arrivals are widely recognized as a tool for modeling transportation networks, flexible manufacturing and telecommunication systems.

Various researchers have considered various models of this type, either to solve particular problems at hand or to develop some results of related stochastic process that arise from them.

We have analyzed a single-server queue with finite buffer and queue length dependent service times. We put N thresholds on buffer. Arrivals are assumed to follow a Compound Poisson Process. The service times depend on queue length at service-start epochs.

There are many articles with studies of queues with length dependent arrival rates and/or service times. The comprehensive surveys of them are given by Dshalalow [1] and Choi [2], [3].

For the analysis we will use the approach suggested by Niu and Cooper in [4] and [5], and we have employed Embedded Markov chain of queue length at service-start epochs and sample path analysis of trajectory of $M^x/G1/K$ queue length.

The rest of this paper is organized as follows: Section 2 consist of model description. In Section 3 we derive some useful preliminary results. An embedded Markov chain at service-start epochs is introduced in Section 4. The queue length and remaining service time distribution is studied by sample path analysis in Section 5.

2. The model description

We consider a $M^x/G/1/K$ queue, where K equals the number of waiting places in the queue, including the space for the customer that may be in service. Arrivals of customers occur according to a Compound Poisson Process. The customers arriving when the buffer is full are lost. When an arrival batch is larger in size than number of available free

waiting places, the free positions will be filled up and the remaining customers of the batch will be lost. We shall introduce the threshold values L_r $(L_{r-1} < L_r, 1 \le r \le N, L_0 = 0, L_n = K)$ on the buffer. If the queue length at service-start epoch of a customer belongs to the interval $[L_{r-1}, L_r)$ the distribution function of service time is $G_r(t)$, mean value $g_{r,1}$. The arrival epochs of the batches form a Poisson process with intensity λ, consecutive batch sizes are independent and have the common probability generating function $C(z) = \sum_{k=1}^{\infty} c_k z^k$ with finite mean number of customers C_1 in each batch.

Customers accepted by the system are served by a single server exhaustively, i.e. the server serves the queue continuously until the queue is empty.

3. The model description

Let us suppose that customers arrive in batch according to time-homogeneous Poisson process at rate $\lambda > 0$. The batch size is a random variable with the probability generating function of $C(z)$.

In fact, the arrivals form a Compound Poisson process so that the probability that i customers arrive in $(0,t]$ is equal to

$$\tilde{a}_i(t) = a_0(t), \text{ if } i = 0 \text{ and } \sum_{k=0}^{i} a_k(t) c_i^{(k)}, \; i = 1, 2, \cdots. \tag{1}$$

where $a_i(t) = \dfrac{(\lambda t)^i e^{-\lambda t}}{i!}$, and $c_i^{(k)}$ is the i-th element of k-th fold convolution of $\{c_i\}$ with itself. The probability generating function of $\tilde{a}_i(t)$ is equal to

$$\sum_{i=0}^{\infty} \tilde{a}_i(t) z^i = \exp\left(-\lambda(1 - C(z))t\right). \tag{2}$$

Then the probability $\tilde{a}_i^{(r)}$ that i customers arrive during a service time with distribution function $G_r(t)$ will equal to

$$\tilde{a}_i^{(r)} = a_0^{(r)} \text{ if } i = 0 \text{ or } \sum_{k=1}^{i} a_k^{(r)} c_i^{(k)}, \text{ if } i = 1, 2, \cdots \tag{3}$$

where $a_k^{(r)} = \int_0^{\infty} \dfrac{(\lambda t)^k e^{-\lambda t}}{k!} dG_r(t)$, and the generating function of $\tilde{a}_i^{(r)}$, $i = 1, 2, \cdots$ and $a_i^{(r)}, i = 1, 2, \cdots$ will equal to $G_r^*(\lambda(1 - C(z)))$ and $G_r^*(\lambda(1-z))$, respectively.

4. The queue length distribution at service-start epochs

Let $\tau_k, (k \geq 0)$ be the successive service-start epochs with $\tau_0 = 0$ For $k \geq 1$ let Q_k be the number of customers waiting in the queue immediately after the k -th service start epoch, and let N_k be the number of customers who arrive during the k -th service interval.

Then, Q is defined to be the process $\{Q_k, k \geq 1\}$, with Q_k determined recursively by

$$Q_{k+1} = \max\{0, Q_k + \min[N_k, K-1-Q_k]-1\}. \quad (4)$$

Note, that N_k is dependent only on Q_k but not on the $Q_l, l \leq k-1$. If $Q_k = j$ and $j \in J_r \equiv [L_{r-1}, L_r)$, then

$$P(N_k = l) = \tilde{a}_l^{(r)}, 0 \leq l < \infty \quad (5)$$

All our results will be expressed in terms of stationary probabilities of Q. The state space of Q is designed to be the set of $\hat{0}, \hat{1}, \hat{2}, \cdots K-\hat{1}$ and they are interpreted as the number of customers in queue after the service start-epoch of the first customer in the busy period, and $0,1,2,\cdots,K-2$ are the queue length after the service-start epoch of any other customer in the busy period. From (4) and (5), it follows that Q is irreducible and aperiodic with finite states and therefore, its stationary probabilities $\hat{\sigma}_0, \hat{\sigma}_1, \hat{\sigma}_{"2}, \cdots, \hat{\sigma}_{K-1}, \sigma_0, \sigma_1, \sigma_2, \cdots, \sigma_{K-2}$ are uniquely determined.

The stationary distribution of the embedded Markov chain is the solution to the next set of linear equations

$$\hat{\sigma}_i = (\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} c_{i+1}, 0 \leq i \leq K-2,$$

$$\hat{\sigma}_{K-1} = (\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} \overline{c}_K, \overline{c}_K = \sum_{l=K}^{\infty} c_l \quad (6)$$

$$\sigma_i = \sum_{r=1}^{r=N} \sum_{l=0}^{l=i+1} (\hat{\sigma}_l + \sigma_l)\hat{a}_{i+1-l}^{(r)} I_r(l), 0 \leq i \leq K-3$$

and $\sum_{i=0}^{K-1} \hat{\sigma}_i + \sum_{i=0}^{K-2} \sigma_i = 1,$

where $I_r(l) = 0$ if $l \notin [L_{r-1}, L_r)$ or $I_r(l) = 1$ if $i \in [Lr-1, L_r)$, and $J_r = [L_{r-1}, L_r)$.

Dividing both sides of (6) by σ_0 and introducing variables $\chi_i = \frac{\sigma_i}{\sigma_0}$, $i = 0,1,2,\cdots,K-2$ and $\hat{\chi}_i = \frac{\hat{\sigma}_i}{\sigma_0}, i = \hat{0}, \hat{1}, \hat{2}, \cdots, K-\hat{1}$ we obtain the recursive relations for $\hat{\chi}_i$ and χ_i as follows

$$\hat{\chi}_0 = \tilde{a}_0^{(1)} c_{i+1} (1 - \tilde{a}_0^{(1)} c_{i+1})^{-1},$$

$$\hat{\chi}_i = (\hat{\chi}_0 + 1)\tilde{a}_0^{(1)} c_{i+1}, 1 \le i \le K-2,$$

$$\hat{\chi}_{K-1} = (\hat{\chi}_0 + 1)\tilde{a}_0^{(1)} \overline{c}_K,$$

$$\chi_0 = 1, \chi_{i+1} = [\chi_i - \sum_{r=1}^{N} \sum_{l=o}^{i} (\hat{\chi}_l + \chi_l)\tilde{a}_{i+1-l}^{(r)} I_r(l) - \sum_{r=1}^{N} \tilde{\chi}_{i+1} . \tilde{a}_0^{(r)} I_r(i+1)] \sum_{r=1}^{N} I_r(i+1)(\tilde{a}_0^{(r)})^{-1}$$

$1 \le i \le K-3.$

From the normalization condition follows

$$\hat{\sigma}_i = \hat{\chi}_i (\sum_{i=0}^{K-1} \hat{\chi}_i + \sum_{i=0}^{K-2} \chi_i)^{-1}, 0 \le i \le K-1, \sigma_i = \chi_i (\sum_{i=0}^{K-1} \hat{\chi}_i + \sum_{i=0}^{K-2} \chi_i)^{-1}, 0 \le i \le K-2.$$

5. The queue length and remaining service time distribution

We will introduce the next notations:

$L(t)$ -the number of customers in the system at time t;

$R(t)$ -when $L(t) > 0$ and $R(t)$ is the remaining service time of the customer in service;

$Q_s(t)$ is the number of customers waiting in queue immediately after the start of the cuurent service;

$Q_a(t)$ -the number of customers arriving from the start of the current service up to time t.

Our goal is to find the steady state distribution of the random process $Z(t)$ where

$Z(t) = 0$ if $L(t) = 0$ or $(L(t), R(t))$, if $L(t) > 0.$ (7)

It is easy to see that $L(t) = 1 + Q_s(t) + \min[Q_a(t), K - 1 - Q_s(t)].$

Then instead of $Z(t)$ we will work with the Markov process $Z^+(t)$ defined as

$Z^+(t) = 0$ if $L(t) = 0$ or $(Q_a(t), Q_s(t), R(t))$ if $L(t) > 0.$ (8)

We will analyze the random process $Z^+(t)$ from the view point of a randomly selected arriving group of customers. We say that, at a given time t the process $Z^+(t)$ is in state θ_0 if $L(t) = 0$, and that it is in state $\theta_{ij}^+(x)$ if $L(t) > 0, Q_s(t) = i, Q_a(t) = j, R(t) \le x$

where $x \ge 0, 0 \le i \le K-2, 0 \le j < \infty,$ and $\hat{\theta}_{ij}^+(x)$ if

$L(t) > 0, Q_s(t) = \hat{i}, Q_a(t) = j, R(t) \le x$, where $x \ge 0, 0 \le \hat{i} \le K-1, 0 \le j < \infty.$

We also assume that $Z^+(A_k)$ is the state of Z^+ as seen by the k-th arriving group of customers, where $\{A_k, k \ge 1\}$ is a consequence of group arrival epochs. We shall study the following proportions associated with $Z^+(t)$:

$$\alpha_0^+ \equiv \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} 1_{\theta_0^+}(Z^+(A_k)) \tag{9}$$

and for $0 \le i \le K-2, 0 \le j < \infty,$ and $x \ge 0$

$$\alpha_{ij}^+(x) \equiv \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} 1_{\theta_{ij}^+(x)}(Z^+(A_k)),$$

$$\hat{\alpha}_{ij}^+(x) \equiv \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} 1_{\hat{\theta}_{ij}^+(x)}(Z^+(A_k)), 0 \le \hat{i} \le K-1 \tag{10}$$

where $1_\theta(.)$ denotes the indicator function of a given $\theta.$ The right hand side of (9) and (10) can be interpreted as the "probability" of the state of Z^+ as seen by a random-selected arriving group of customers. At first, we shall determine α_0^+. The rate at which customers arrive into the system is equal to λC_1 and they brought into the system the work equals to $\lambda C_1 \sum_{r=1}^{N} g_{r,1} \sum_{j\in J_r} I_r(j)(\hat{\sigma}_j + \sigma_j)$. But only a part of them are entered and will be served. The proportion of entering customers, by definition is equal to η. Then the customers who are accepted by the server and will be served, brought into the system work equal to

$$\eta\ \lambda C_1 \sum_{r=1}^{N} g_{r,1} \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j) = \eta\ C_1 \sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j), \text{ where } \rho_r = \lambda g_{r,1}.$$

On the other hand the proportion of time during which the server is busy is equal to

$$\frac{\sum_{r=1}^{N} g_{r,1} \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)}\frac{1}{\lambda} + \sum_{r=1}^{N} g_{r,1} \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)} = \frac{\sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)}$$

and since Poisson arrival see time averages also equals to the proportion of arriving group of customers that are blocked on their arrival.

Thus we have

$$1-\alpha_0^+ = \eta C_1 \sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j) = \frac{\sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)} \tag{11}$$

Then

$$\alpha_0^+ = \frac{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)}}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j\in J_r} (\hat{\sigma}_j + \sigma_j)}$$

$$\eta = \frac{1}{C_1[(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)]} \tag{12}$$

To obtain the probability $\alpha_{ij}^{+}(x)$ we shall consider the remaining variables $Q_a(\cdot)$ and $R(\cdot)$ in Z^{+}, but now from the viewpoint of a randomly-selected blocked group of customers. Let $\theta_{i\cdot}^{+}(\infty)$, $0 \le i \le K-2$ ($\hat{\theta}_{i\cdot}^{+}(\infty), 0 \le \hat{i} \le K-1$) denote the state of Z^{+}, if $L(t) > 0, (Q_s(t) = i),$ and $1_{i\cdot}$ ($1_{\hat{i}\cdot}$) will be the indicator function of the event that Z^{+} is in state $\theta_{i\cdot}^{+}(\infty)$, ($\hat{\theta}_{i\cdot}^{+}(\infty)$) immediately after the k-th service-start epoch. We still denote by $1_{\cdot j/i;x}(k)$, ($1_{\cdot j/\hat{i};x}(k)$) the indicator function of the event that, of the N_k group of customers (all of which are blocked) that arrived during the k-th service time, there is one group, which on arrival find in the queue j customers arrived since the service-start epoch of the same k-th service time and $Q_s(t) = i$ $(Q_s(t) = \hat{i})$. Let us define functions

$$\beta_{ij}^{+}(x) \equiv \frac{\alpha_{ij}^{+}(x)}{1-\alpha_0^{+}}, \quad \beta_{ij}^{+}(\infty) \equiv \frac{\alpha_{i\cdot}^{+}(\infty)}{1-\alpha_0^{+}}, \text{for } 0 \le j,\ 0 \le i \le K-2 \tag{13}$$

and $\hat{\beta}_{ij}^{+}(x) \equiv \frac{\hat{\alpha}_{ij}^{+}(x)}{1-\alpha_0^{+}}$, $\hat{\beta}_{ij}^{+}(\infty) \equiv \frac{\hat{\alpha}_{i\cdot}^{+}(\infty)}{1-\alpha_0^{+}}$, for $0 \le i \le K-1, j \ge 0, x \ge 0.$ (14)

Since $(1-\alpha_0^{+})$ is the proportion of arriving groups that are blocked, and $\alpha_{ij}^{+}(x)$ is the proportion of arriving groups that find the system in state $\theta_{ij}^{+}(x)$, we can take the view point of a randomly selected blocked group of customers and rewrite $\beta_{ij}^{+}(x)$ in the following equivalent, explicit form: $\beta_{ij}^{+}(x) = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} 1_{ij;x}(k)}{\sum_{k=1}^{n} N_k}$ where N_k is the number of of arriving groups and $\beta_{ij}^{+}(\infty)$ is the proportion of arriving group of customers that are blocked and find a system at state $\theta_{i\cdot}^{+}(\infty)$. Then $\beta_{i\cdot}^{+}(\infty) = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} 1_{i\cdot}(k)}{\sum_{k=1}^{n} N_k}$. We shall interpret the number of group of customers who arrive during a service interval that start with $Q_s(\cdot) = i\cdot$ as "a sojourn in the state i" in discrete time Semi-Markov process. Then it is easily seen that $\beta_{i\cdot}^{+}(\infty)$ is the proportion equal to

$$\beta_{i\cdot}^{+}(\infty)=\frac{\sigma_i\sum_{r=1}^{N}\lambda g_{r,1}I_r(i)}{\sum_{i=0}^{K-1}\hat{\sigma}_i\sum_{r=1}^{N}\lambda g_{r,1}I_r(i)+\sum_{i=0}^{K-2}\sigma_i\sum_{r=1}^{N}\lambda g_{r,1}I_r(i)}$$

$$\frac{\sigma_i\sum_{r=1}^{N}g_{r,1}I_r(i)}{\sum_{i=0}^{K-1}\hat{\sigma}_i\sum_{r=1}^{N}g_{r,1}I_r(i)+\sum_{i=0}^{K-2}\sigma_i\sum_{r=1}^{N}g_{r,1}I_r(i)} \tag{15}$$

Now we shall consider

$$\frac{\beta_{ij}^{+}(x)}{\beta_i^{+}(\infty)}=\lim_{n\to\infty}\frac{\sum_{k=1}^{n}1_{ij;x}(k)/\sum_{k=1}^{n}N_k}{\sum_{k=1}^{n}1_{i\cdot}(k)N_k/\sum_{k=1}^{n}N_k}=\lim_{n\to\infty}\frac{\sum_{k=1}^{n}1_{ij;x}(k)}{\sum_{k=1}^{n}1_{i\cdot}(k)N_k}$$

Since $1_{ij;x}(k)\equiv 1_{i\cdot}(k)1_{\cdot\, j/i;x}(k)$, then from $1_{i\cdot}(k)=0$ in both the denominator and numerator in the last expression, it is further simplified to $\frac{\beta_{ij}^{+}(x)}{\beta_i^{+}(\infty)}=\lim_{m\to\infty}\frac{\sum_{l=1}^{m}1_{ij;x}(k_l)}{\sum_{l=1}^{m}1_{i\cdot}(k_l)N_{k_l}}=\lim_{m\to\infty}\frac{\sum_{k=1}^{n}1_{ij;x}(k_l)}{\sum_{l=1}^{m}N_{k_l}}$, where $\{k_l,\}\ \{k_l, l=1,2,\cdots\}$ enumerates the subset of the indices $k=1,2,\cdots$ for which $1_{i\cdot}(k)=1.\ 1_{i\cdot}(k)=1.$

Finally, since $\{1_{\cdot\, j/i;x}(k), k\geq 1\}$ are sequences of i.i.d. random variables from SLLN follows

$$\frac{\beta_{ij}^{+}(x)}{\beta_i^{+}(\infty)}=\lim_{m\to\infty}\frac{\sum_{l=1}^{m}1_{\cdot\, j/i;x}(k_l)/m}{\sum_{l=1}^{m}N_{k_l}/m}=\frac{E[1_{\cdot\, j/i;x}(1)]}{\sum_{r=1}^{N}\lambda g_{r,1}I_r(i)}.$$

Conditioning on the time of arrival after service-start epoch, we have

$$E[1_{\cdot\, j/i;x}(1)]=P(1_{\cdot\, j/i;x}(1)=1)=\sum_{r=1}^{N}I_r(i)\int_0^{\infty}[G_r(y+x)-G_r(y)]\tilde{a}_j(y)\lambda dy.$$

Then $\frac{\beta_{ij}^{+}(x)}{\beta_{i\cdot}(\infty)}=\frac{\sum_{r=1}^{N}I_r(i)\int_0^{\infty}[G_r(y+x)-G_r(y)]\tilde{a}_j(y)\lambda dy}{\sum_{r=1}^{N}I_r(i)\lambda g_{r,1}}\equiv\nu_{ij}(x)$ (16)

Analogous considerations lead to $\hat{\beta}_{i\cdot}^{+}(\infty) = \dfrac{\hat{\sigma}_i \sum_{r=1}^{N} g_{r,1} I_r(i)}{\sum_{i=0}^{K-1} \hat{\sigma}_i \sum_{r=1}^{N} g_{r,1} I_r(i) + \sum_{i=0}^{K-2} \sigma_i \sum_{r=1}^{N} g_{r,1} I_r(i)}$ (17)

$$\frac{\hat{\beta}_{ij}^{+}(x)}{\hat{\beta}_{i\cdot}^{+}(\infty)} = \frac{\sum_{r=1}^{N} I_r(i) \int_0^{\infty} [G_r(y+x) - G_r(y)] \tilde{a}_j(y) dy}{\sum_{r=1}^{N} I_r(i) g_{r,1}} \equiv \nu_{ij}(x) \qquad (18)$$

Thus we shall formulate our main result in Theorem 1.

Theorem 1. The state distribution of Z^+ as seen by a randomly-selected arriving group of customers is, w. p. 1, given by

$$\alpha_0^{+} = \frac{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)}}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)}, \alpha_{ij}^{+}(x) = (1 - \alpha_0^{+}) \beta_{i\cdot}(\infty) \nu_{ij}(x) =$$

$$\frac{\sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)} \frac{\sigma_i \sum_{r=1}^{n} g_{r,1} I_r(i)}{\sum_{i=0}^{K-1} \hat{\sigma}_i \sum_{r=1}^{N} g_{r,1} I_r(i) + \sum_{i=0}^{K-2} \sigma_i \sum_{r=1}^{N} g_{r,1} I_r(i)} \nu_{ij}(x),$$

$$\hat{\alpha}_{ij}^{(x)}(x) = \frac{\sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)}$$

$$\frac{\hat{\sigma}_i \sum_{r=1}^{n} g_{r,1} I_r(i)}{\sum_{i=0}^{K-1} \hat{\sigma}_i \sum_{r=1}^{N} g_{r,1} I_r(i) + \sum_{i=0}^{K-2} \sigma_i \sum_{r=1}^{N} g_{r,1} I_r(i)} \nu_{ij}(x).$$

The proportion η of arriving customers who actually enter in the system and receive service w. p. 1 is equal to $\eta = \dfrac{1}{C_1[(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)]}$.

Now, let us consider the process Z. We say that Z, at a given time t, is in state θ_0 if $L(t) = 0$, and that it is in state $\theta_j(x)$, where $1 \le j \le K$ and $x \ge 0$, if $L(t) = j$ and

$R(t) \le x$. We define α_0 and $\alpha_j(x), 1 \le j \le K$ and $x \ge 0$, as the proportions of group of customers including those who do not enter or who on arrival find Z in state θ_0 and $\theta_j(x)$, respectively. Then, from Theorem 1 follows Theorem 2.

Theorem 2 The state distribution of Z as seen by a randomly-selected arriving group of customers is, w.p.1, given by

$$\alpha_0 = \alpha_0^+,$$

$$\alpha_j(x) = \sum_{i=0}^{j-1} (\hat{\alpha}_{i,ji-1}(x) + \alpha^+_{i,j-i-1}(x)),\ 1 \le j \le K-1,\ x \ge 0$$

$$\alpha_K(x) = \sum_{j=K}^{\infty} \sum_{i=0}^{K-2} \alpha^+_{i,j-i-1}(x) + \sum_{j=K}^{\infty} \sum_{i=0}^{K-1} \hat{\alpha}^+_{i,j-i-1}(x),\ \ x \ge 0\,.$$

Theorem 3. The stationary distribution of the number of customers in the system is given by

$$\alpha_0 = \frac{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)}}{(\hat{\sigma}_0 + \sigma_0)\tilde{a}_0^{(1)} + \sum_{r=1}^{N} \rho_r \sum_{j \in J_r} (\hat{\sigma}_j + \sigma_j)},$$

$$\alpha_j(\infty) = \sum_{i=0}^{j-1} (\hat{\alpha}_{i,ji-1}(\infty) + \alpha^+_{i,j-i-1}(\infty)),\ 1 \le j \le K-1,$$

$$\alpha_K = \alpha_K(\infty) = \sum_{j=K}^{\infty} \sum_{i=0}^{K-2} \alpha^+_{i,j-i-1}(\infty) + \sum_{j=K}^{\infty} \sum_{i=0}^{K-1} \hat{\alpha}^+_{i,j-i-1}(\infty) = 1 - \sum_{j=0}^{K-1} \alpha_j.\ \ x \ge 0\,.$$

The probability that the first customer of a batch is lost is given by α_K.

Proof: Because of the arrival process of batches is Poisson and Poisson Arrivals See Time Averages Property, then the first customer of a batch see time averages and it will see all waiting place busy with the probability α_K and the stationary distribution of queue length is α_0, and $\alpha_j(\infty)$, $1 \le j \le K$.

References

1. J.H. Dshalalow, Queueing systems with state dependent parameters, In: Frontiers in Queueing Models with Apppl. in Science and Eng, (ed . by J.H. Dshalalow), CRC Press, Boca Ratton, 1997, 61-116.
2. B. D. Choi, Y. C. Kim, Y. W. Shin, C. E. M. Pearce , The $M^x / G/1$ queue with queue length dependent service times, Appl. Math. and Stoch. Anal., 2001, 14, 399-419.
3. D.I. Choi, T.S. KIM, S. Lee, Analysis of an $MMPP/GI/1/K$ queue with queue length dependent arrival rates, and its application to preventive congestion control in telecommunication networks, European j. Oper. Res., 2008, 187, 652-659.
4. S.-C. Niu, Representing Workloads in $GI/G/1$ queues through the Preemptive-Resume LIFO queue discipline, Queueing Systems, 1988, 3, 157-178.
5. S.-C. Niu, R. B. Cooper, Transform-free analysis of $M/G/1/K$ and related queues, Mathematics of Operations Research, 1993, 18, 2, 486-510.

Orthogonal Arrays and Software Testing

Petya Valcheva

UNWE, 1700 Sofia, Bulgaria
pvalcheva@unwe.bg

Abstract. The software development has become an intrinsic part of our life in recent years. As a result, many prosperous companies realize that software testing is an investment in quality. Orthogonal arrays are combinatorial structures widely spread in many scientific fields. One of the most important applications in computer science is the Orthogonal Array Testing System (OATS). It is a mathematically based algorithm, which efficiently optimizes the number of tests for new software products. This article deals with the mathematical argumentation and the application of this method, which provides maximum coverage with nine test cases, when we have three new processors to be tested on a networked computer, specified by his three components.

Keywords. Orthogonal array testing system, software testing, design of experiment, Latin square, difference matrix.

1. Introduction and History

The latest researches indicates that a huge amounts of money are spent on researching new efficient ways and tools to cope with minimizing testing time and cost. A widely cited report of National Institute of Standards and Technology (NIST) from 2002 indicates that more than 50 percent of software development budgets go to testing, but on the other hand omissions still cost the U.S. economy $59.5 billion annually [1]. These data clearly show the urgent need for innovative methods in which even small improvements in testing approach would give positive results in cost savings.

A software tester have to attempt to break the system in all possible ways, so that to be able to detect all bugs, which will therefore provide qualitative software to the users. Orthogonal Array Testing System is an appropriate statistical way of testing pairwise interactions by deriving suitable and compact set of test cases.

The evolution of the method has in fact long and dynamic progress through the years. Design of experiment was first invented by Ronald A. Fisher in the 1920s at Rothamsted Experimental Station for an agricultural research [2]. He showed how different conclusions could be done effectively from experiments with natural fluctuations such as temperature, soil and fertilizers. Design of Experiment is effective tool applicable to many areas such as physics, statistics, computer simulations models and others. It is a series of tests, in which are made purposeful changes to the input variables in order to investigate the effects of many different factors by varying them simultaneously. As a result of researching work in the Combinatorics based on the concepts of Galois fields and finite projective geometrics, Rao defined the combinatorial configurations Orthogonal arrays in 1943 [3]. In the last sixty years orthogonal arrays have been actively used in the Design of experiment in agricultural and medical tests, marketing, and software industry too. Genichi Taguchi began to work in 1950 actively on improve the quality and reliability in production and his methods are widely used today [4]. In 1980 started the work upon implementation of the ideas in experimental

design applied to software testing. Keizo Tatsumi from Fujitsu first described the usage of orthogonal and covering arrays for generating test configurations. Later Robert Mandl from Softech used orthogonal Latin squares to design tests for compilers.

The article below aims to demonstrate mathematical reasoning and results of applying the method in a concrete example. It will have less parameters in order to show better the mathematical algorithm and the conclusions done by the application of the concrete way. It will be defined four test factors, each of which may vary in three values. In terms of combinatorics this would determine 81 possible configurations. In the next part it will be shown how to construct appropriate orthogonal array according to the input parameters. At the end of the article the problem will be decided by Orthogonal Array Testing System, where will clearly be demonstrated the significant reduction of configurations in test set of 9 cases. But, of course, the real systems are more complicated. For example, if it is needed to test system with 6 factors, each of which with 5 different components, the number of possible configurations will grow to $5^6 = 15625$, but Orthogonal Array Testing System will reduce them to 25 test cases.

2. On the Constructions of Orthogonal Arrays

The purpose of this section is to acquaint the reader with the basic ideas and the development of a comprehensive toolkit for the construction of orthogonal arrays (OAs) as a means of application in software interaction testing. With the aim of better understanding we will stop our attention to the most commonly used cases described below.

Orthogonal array $OA(n,k)$ is a $k \times n^2$ matrix with entries from a set N of cardinality n, usually $N = \{0,1,\dots,n-1\}$ (residues modulo n), such that any two rows contain all the ordered pairs from the set N. Since the number of these pairs is n^2, it is equal to the number of columns in OA and consequently every pair appears exactly once. It is customary to say that the orthogonal array has *index unity* and *strength* k. It is well known that $k \leq n+1$.

For example, the following table presents an $OA(3,4)$

$$OA(3,4) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 2 & 1 & 1 & 0 & 2 & 2 & 1 & 0 \end{array}\right)$$

Table 1: $OA(3,4)$

Orthogonal arrays are often constructed via algebraic arguments. The most common method of deriving such array is through the *difference matrix* $DM(m,n)$. A full description of various types DMs and their properties can be found in [5], [6] and [7]. Let G be an abelian (not necessary) group of order n. A difference matrix $DM(m,n)$ is defined as a $m \times n$ matrix $A = (a_{ij})\ i = 1,\dots,k\ ,j = 1,\dots,n, a_{ij} \in G$, any two rows of which give all elements of G, presented as differences, i.e. for arbitrary $1 \leq i_1 \neq i_2 \leq m$ the set $\{a_{i_1 j} - a_{i_2 j}, j = 1,\dots,n\}$ coincide with G. The orthogonal array $OA(n,m)$ is now generated by the following construction:

(1) $$OA = (A \mid A+1 \mid \dots \mid A+(n-1))$$

where $A+X$ is the matrix $A_i = (a_{ij} + X)$.The array in $Table\ 1$, $OA(3,4)$, can be obtained as follows: First, we give difference matrix $DM(3,3)$

(2) $$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

Then, after applying the aforementioned construction in (1), we will reach to the following orthogonal array $OA(3,3)$

$$OA(3,3) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\ \underbrace{0}_{A} & \underbrace{2}_{} & \underbrace{1}_{} & \underbrace{1}_{A+1} & \underbrace{0}_{} & \underbrace{2}_{} & \underbrace{2}_{A+2} & \underbrace{1}_{} & \underbrace{0}_{} \end{array}\right)$$

The last array can be extended up to $OA(3,4)$ by adding a row in a special way.

A *transversal* in $OA(n,k)$ is a $k \times n$ sub-array, any row of which is a permutation of the elements of G, i.e. the transversal is a set of n columns of OA with no common elements. A set of n columns-disjoint transversals- is called a $resolution$ of OA. In the concrete example, the resolution we need consists of 3 transversals:

Transversal	Columns	Element
1	1,4,7	0
2	2,5,8	1
3	3,6,9	2

Table 2: Transversals of $OA(3,3)$

Adding to every column of a transversal an element from G one obtains row 4 of $OA(3,4)$. In details: add 0 to columns 1,4,7; 1 to columns 2,5,8 and 2 to columns 3,6,9 and will obtain following array from *Table 1*

$$OA(3,4) = \left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 2 & 1 & 1 & 0 & 2 & 2 & 1 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{0} & \mathbf{1} & \mathbf{2} \end{array}\right)$$

When we interchange the rows of the above array, we will reach to the $OA(3,4)$ of $Table\ 1$.

Another approach to orthogonal arrays are constructions derived from $Latin\ squares$ (LS), a technique introduced by Leonhard Euler. A Latin square of order n is an $n \times n$ array with entries from a set N, such that each element of N appears once in every row and every column(i.e. every row and every column is a permutation of the elements of N).

Returning to extended Table 3, the LS can be obtained from OA in the following manner: the first row in $OA(3,4)$ from $Table\ 1$ presents the number of row in the Latin square, the second row gives the number of column, and the third and fourth - the entries in the squares. Consequently, row 3 and row 4 give two LS of order 3, shown below:

$$\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{vmatrix} \qquad \begin{vmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ 2 & 1 & 0 \end{vmatrix}$$

Denoting the elements of the first square with a_{ij} and of the second $-b_{ij}$, the pairs $(a_{ij}, b_{ij})\ i,j = 1,2,3$ are all the ordered pairs of 3 elements $\{0,1,2\}$. Such squares are called $orthogonal$. Generally, two Latin squares $LS_1 = (a_{ij})$ and $LS_2 = (b_{ij})$ of order n are called orthogonal if the pairs $(a_{ij}, b_{ij})\ i,j = 1,\dots,n$ are all the ordered pairs of n elements.

Several Latin squares are called *mutually orthogonal* ($MOLS$), if any two of them are orthogonal.

Theorem: The existence of $OA(n,k)$ is equivalent to the existence of $k-2$ $MOLS$ of order n ▫

Returning to the strength k of $OA(n,k)$ no more than $n-1$ $MOLS(n)$ do exist.

Above construction shows how to construct the $OA(n,k)$, but we need construction of $DM(k-1,n)$ (note that DM is n times shorter than the OA itself). Difference matrices are often constructed via another algebraic structures called *Galois fields* - a field of finite number of elements. The residues of prime module form field through addition and multiplication of the residues. For example, the field of 3 elements $GF(3)$ is defined by two tables for addition and multiplication

+	**0**	**1**	**2**
0	0	1	2
1	1	2	0
2	2	0	1

*	**0**	**1**	**2**
0	0	0	0
1	0	1	2
2	0	2	1

The third row of the difference matrix A can be derived as multiplication of the second row by 2. More generally, given field $GF(p)$, p is a prime, take two rows

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & p-1 \end{pmatrix}$$

and multiply the second of them by the non-zero elements of $GF(p)$. So the resulting matrix is

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & p-1 \\ 0 & 2 & \cdots & 2(p-1) \\ \vdots & \vdots & & \vdots \\ 0 & (p-1) & \cdots & (p-1)^2 \end{pmatrix}$$

which is (as can be easily seen) $DM(p,p)$. According to the above, this matrix can be extended up to $OA(p,p+1)$

Example for $OA(5,6)$

As we mention at the beginning the real systems contain more factors. As an example we will construct $OA(5,6)$, as first construct $DM(5,5)$ over $GF(5)$ and then extend to $OA(5,5)$, after that by adding a row will reach to the $OA(5,6)$.

The Galois field is $GF(5) = \{0,1,2,3,4\}$. The corresponding difference matrix $DM(5,5)$ over this field is

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \\ 0 & 3 & 1 & 4 & 2 \\ 0 & 4 & 3 & 2 & 1 \end{pmatrix}$$

The new orthogonal array $OA(5,5)$ will be derived after applying the rule (1) with the matrix B. The resulting orthogonal array consists of five blocks, because $n = 5$. Consequently, $OA(5,5)$ is of the form:

$$OA(5,5) = \left(\begin{array}{ccccc|ccccc|ccccc|ccccc|ccccc} 0&0&0&0&0&1&1&1&1&1&2&2&2&2&2&3&3&3&3&3&4&4&4&4&4\\ 0&1&2&3&4&1&2&3&4&0&2&3&4&0&1&3&4&0&1&2&4&0&1&2&3\\ 0&2&4&1&3&1&3&0&2&4&2&4&1&3&0&3&0&2&4&1&4&1&3&0&2\\ 0&3&1&4&2&1&4&2&0&3&2&0&3&1&4&3&1&4&2&0&4&2&0&3&1\\ 0&4&3&2&1&1&0&4&3&2&2&1&0&4&3&3&2&1&0&4&4&3&2&1&0 \end{array}\right)$$

The corresponding transversals are shown in the Table:

Transversal	Columns	Element
1	1,6,11,16,21	0
2	2,7,12,17,22	1
3	3,8,13,18,23	2
4	4,9,14,19,24	3
5	5,10,15,20,25	4

Table 3: Transversals of $OA(5,5)$

The orthogonal array $OA(5,6)$ will be derived from the last array $OA(5,5)$, after adding an additional row using the elements of the group $GF(5)$ and still founded transversals in *Table 3* as follows: put the element 0 in the last row of the first transversal, i.e. in the columns with numbers 1, 6, 11, 16 and 21. Analogously, for the rest elements from $GF(5)$ and another four transversals.

$$OA(5,6) = \left(\begin{array}{ccccc|ccccc|ccccc|ccccc|ccccc} 0&0&0&0&0&1&1&1&1&1&2&2&2&2&2&3&3&3&3&3&4&4&4&4&4\\ 0&1&2&3&4&1&2&3&4&0&2&3&4&0&1&3&4&0&1&2&4&0&1&2&3\\ 0&2&4&1&3&1&3&0&2&4&2&4&1&3&0&3&0&2&4&1&4&1&3&0&2\\ 0&3&1&4&4&1&4&2&0&3&2&0&3&1&4&3&1&4&2&0&4&2&0&3&1\\ 0&4&3&2&1&1&0&4&3&2&2&1&0&4&3&3&2&1&0&4&4&3&2&1&0\\ \mathbf{0}&\mathbf{1}&\mathbf{2}&\mathbf{3}&\mathbf{4}&\mathbf{0}&\mathbf{1}&\mathbf{2}&\mathbf{3}&\mathbf{4}&\mathbf{0}&\mathbf{1}&\mathbf{2}&\mathbf{3}&\mathbf{4}&\mathbf{0}&\mathbf{1}&\mathbf{2}&\mathbf{3}&\mathbf{4}&\mathbf{0}&\mathbf{1}&\mathbf{2}&\mathbf{3}&\mathbf{4} \end{array}\right)$$

Table 4: $OA(5,6)$

3. An Example for Component Based System

The selection of test cases is one of the biggest challenges for the Quality Assurance (QA) engineer. It is known that the testing process can only show the existence defects and never their absence. For that reason, it is better to choose executing a concise, well-defined set of tests that are likely to uncover most (not all) of the bugs in comparison to endlessly executing tests that just aren't likely to find mistakes and don't increase your confidence in the system.

The ultimate aim of the section presented below is to demonstrate the practical application of the previously constructed design $OA(3,4)$.

A company manufactures processors for the PC market. They are made from pre-built and pre-tested hardware components. Different processors are designed to work on a variety of commercial operating systems and to support several kinds of hard disk interfaces.

To be the product accessible, the company allows the system to be sold with different amounts of memory, as shown in the table below:

	C O M P O N E N T S			
	Factor 1	Factor 2	Factor 3	Factor 4
	Processor	Operation System(OS)	Memory Configuration	Disk Interface
	Intel Core i3 Intel Pentium Intel Core i5	Windows 7 Linux Novell Netware 6.x	2GB 8GB 6GB	Ultra 320-SCSI Ultra 160-SCSI Ultra 160-SATA

This table consists of *four* factors (Intel processor, operating system, memory configuration and hard disk interface), each of these has *three* supported values. Taking into account these parameters, it is appropriate to use orthogonal array $OA(3,4)$ from $Table\ 1$. The array will be completed as follows: for the different kind of processors, we will use the different elements from the first row (0 for Intel Core i3, 1 for Intel Pentium and 2 for Intel Core i5). The other three rows will be filled with the corresponding elements in the same way. After inflict all the elements from the table in the orthogonal array, we derive appropriate and compact set of nine test cases.

Intel Core i3	Intel Core i3	Intel Core i3	Intel Pentium	Intel Pentium	Intel Pentium	Intel Core i5	Intel Core i5	Intel Core i5
Windows 7	Linux	Novell Netware 6.x	Windows 7	Linux	Novell Netware 6.x	Windows 7	Linux	Novell Netware 6.x
2GB	8GB	6GB	8GB	6GB	2GB	6GB	2GB	8GB
Ultra 320-SCSI	Ultra160-SCSI	Ultra 160-SATA	Ultra 160-SATA	Ultra320-SCSI	Ultra 160-SCSI	Ultra 160-SCSI	Ultra160-SATA	Ultra 320-SCSI

Each of the nine test configurations in the table above has 4 components with one value from each component selected. The first test configuration (Intel Core i3, Windows 7, 2GB, Ultra 320 –SCSI) covers six *pairs* of interactions (Intel Core i3 with Windows 7, Intel Core i3 with 2GB of memory, Intel Core i3 with an Ultra 320- SCSI disk interface, Windows 7 with 2GB of memory, Windows 7 with an Ultra 320- SCSI disk interface and 2GB of memory with an Ultra 320- SCSI disk interface) or four *triples* of interactions (Intel Core i3 with 2GB of memory and Windows 7, Intel Core i3 with 2GB and an Ultra 320- SCSI disk interface, Windows 7 with 2GB of memory and an Ultra 320- SCSI disk interface, Intel Core i3 with Windows 7 and Ultra 320- SCSI disk interface) [8].

4. Conclusion

One of the advantages of the Pair-wise Testing (or OATS) is the protection of pair bugs. At the same time this test dramatically reduces the number of tests, which is actually the most important, because the pair bugs represent a plurality of combinations bugs that occur much more frequently than those with more variables.

Orthogonal Array Testing System covers 100 %, i.e. nine of nine of the pair-wise combinations, 33 %, i.e. nine of twenty-seven of the three-way combinations and 11 %, i.e. nine of eighty-one of the four-way combinations.

References

1. Kanglin Li, Mengqi Wu, Effective Software Test Automation: Developing an Automated Software Testing Tool, 2004.
2. R. A. Fisher, The Design of Experiments, Hafner Peublishing Company, 1935.
3. A.S. Hedayat, N.J.A Sloane , John Stufken, Ortogonal Arrays: Theory and Applications, Springer-Verlag New York, 1999.
4. Jayaswal, Bijay K, Peter C. Patton , Design for Trustworthy Software: Tools, Techniques, and Methodology of Developing Robust Software, Prentice Hall, 2006.
5. C. Colbourn and J. Dinitz (Editors), CRC Handbook of Combinatorial Designs, CRC Press, Boca Raton, 2007
6. T. Beth, D. Jungnickel, and H. Lenz, Design Theory, Cambridge University Press, Cambridge, England, 1999
7. D. T. Todorov, Four mutually orthogonal Latin squares of order 14, Journal of Combinatorial Designs , p.1-5, 2012
8. Cohen B. Myra, Designing Test Suites for Software Interaction Testing, The University of Auckland, Phd Thesis, 2004

Mathematical Modeling of the Migration of Human Populations

Elena V. Nikolova[1], Ivan P. Jordanov[1,2], Nikolay K. Vitanov[1]

[1] Institute of Mechanics - BAS, Acad. G. Bonchev Str., Bl. 4, 1113 Sofia, Bulgaria

[2] UNWE, 1700 Sofia, Bulgaria
i_jordanov@email.bg

Abstract The dynamics of many social and natural systems is nonlinear. In this aspect we model the spatial and temporal dynamics of human populations by a system of nonlinear partial differential equations. We are interested in the waves caused by immigration of people. We assume that the immigration is a diffusion process influenced by the changing values of the growth rates and coefficients of interaction among the migration groups. The agent models are an important tool for the analysis of complex systems. Depending on the nature of the system, the agents may have a variety of properties, as well as to interact in a different way. In recent years there has been rapid development of the agent models because with their help we can adequately describe the processes in a number of social systems, such as the migration of various separate groups of human population.

Keywords: nonlinear dynamics, migration of human populations, agent models, population dynamics, Dimitrova – Vitanov model, agent systems.

Introduction

The permanent migration of people from one region to another one (within the borders of the country or abroad) is a global problem. In principle, it appears as a result of any economic, political, ideological and social restrictions in the specific country. For each country demographic statistics is available for this migration, but it is not sufficient to predict the further spatio-temporal features in the human population dynamics. Mathematical modeling provides for this possibility and helps to generate testable prognoses. It ensures also a way to design and evaluate protocols, and thereby to manage and control the immigration dynamics. The systems studied as population dynamics are one of the best examples of nonlinear systems in quantitative social sciences [1,2,3,23,24,27]. Nonlinearities arise in the model equations of such systems due to interactions among the individuals or the populations, and the limitations in the environment [14,25]. Such nonlinear model systems require applying methods of nonlinear dynamics [4,8,16], chaos theory [9,13,17], and theory of stochastic processes [13,16,29].

The agent models are an important tool for the analysis of complex systems [5,6,7,8]. Depending on the nature of the system, the agents may have a variety of properties, as well as to interact in a different way. In recent years there has been rapid development of agent models because with their help we can adequately describe the processes in a number of social systems, such as the migration of various separate groups of human population[15,17,25,28].

If the number of agents and chaos in their behavior is small, it is possible to describe the system of interacting agents with more or less - less complicated systems of ordinary or partial differential equations. Such systems should describe the characteristic of the collective behavior of agents (their movement in space). We need just such modeling is required to describe the interactions between different social groups. Therefore, the deterministic model using the various kinds of systems, differential equations is widely spread further. For the case we are interested in pro and spatial detection should be available to individuals in the system of agents we need to use more partial differential equations and systems.

There are different methods for the preparation of a solution of these systems of equations. Some of them are obtained values of the solution corresponding to the specified time and space coordinates. In some cases, however, a good idea will give us the decision itself, although derived under appropriate assumptions about its type. Lately particularly interesting from a practical point of view, are the solutions obtained in the form of waves that describes the immigration of individuals in space, i.e. so-called population waves [18,19,20].

1. Mathematical model of migration

Let us now consider systems of interacting agents, the density of which depend not only on time but also on the coordinates in two-dimensional space. The presence of the spacer dimensions and allows us to study migration. Therefore, the mathematical model is to be a system of partial differential equations. If you solve this system of equations, we can better understand migration processes in social systems.

Let's assume a two-dimensional region D (for example, part of the ground), which are n number of interacting systems of agents, each of which comprises N_i $(i = 1,...,n)$ agents. Then, if we take an arbitrary sufficiently small platform facing dS around any fixed point $\vec{r}$ points of the area in which we dN_i number of individuals of the i-th agent system, we can determine the density of the system agents in the section:

$$\rho_i(\vec{r},t) = \frac{dN_i}{dS}. \tag{1}$$

Each of these density changes with the passage of time. This change is due to the emergence of agents in the area of the site and the entry of agents (i.e. migration) in the boundary of the site.

Let's now in D is a function $R_i(\vec{r},t)$, which describes the density of the agents of the i-th population in that area. Then the time dt in the face dS, i.e. in the point $\vec{r}$ is

$$dN_i^{vol.} = R_i(t,\vec{r})dSdt, \tag{2}$$

from the i-th system agents.

Let us now take an arbitrary line length dl, start at the point $\vec{r}$ and a single normal vector $\vec{n}$. Let time dt pass through the line $I_i(\vec{n})dldt$ by individual i-th system agents. Then, to characterize the migration of individuals at the point $\vec{r}$ we need to know scalar function of the vector argument $I_i(\vec{n})$. If we consider that this function is linear, we can define it by using the vector I_i. This vector we call vector migration.

Indeed let us take a Cartesian coordinate system with origin at the point $\vec{r}$, unit vectors $\vec{i}$ and $\vec{j}$ and a triangle of Cauchy. Then this triangle will appear:

$$dN_i = dN_i^{vol.} + dN_i^{migr.} \tag{3}$$

of subjects in the of *i*-th system agents such as

$$dN_i^{migr.} = I_i(\vec{j})dl_x + I_i(\vec{i})dl_y + I_i(-\vec{n})dl. \tag{4}$$

Above is denoted by dl_x and dl_y are respectively the lengths of the sides of normal $\vec{j}$ and $\vec{i}$, and dl is the length of the sloped side. Let now l is the characteristic size of the triangle then:

$$dN_{i,}dN_i^{vol.} \sim l^2, dN_i^{migr.} \sim l, \tag{5}$$

And if *l* is an infinitesimal quantity, we obtain the equation

$$I_i(\vec{j})dl_x + I_i(\vec{i})dl_y + I_i(-\vec{n})dl = 0. \tag{6}$$

Above is denoted by $\vec{n}$ is denoted with the normal to the inclined side. Note that it is obvious that the $I_i(-\vec{n}) = -I_i(\vec{n}).$ Then above equation we get:

$$I_i(\vec{j})dl_x + I_i(\vec{i})dl_y = I_i(\vec{n})dl. \tag{7}$$

From the above, we get:

$$I_i(\vec{n})dl = I_i(\vec{j})\frac{dl_x}{dl} + I_i(\vec{i})\frac{dl_y}{dl}. \tag{8}$$

Now, you notice that:

$$\begin{aligned} &\frac{dl_x}{dl} = \cos(\vec{i},\vec{n}) = \vec{n}\cdot\vec{i}, \\ &\frac{dl_y}{dl} = \cos(\vec{j},\vec{n}) = \vec{n}\cdot\vec{j}. \end{aligned} \tag{9}$$

We get:

$$I_i(\vec{n}) = \vec{n}\cdot\left[I_i(\vec{j}).\vec{i} + I_i(\vec{i}).\vec{j}\right] = \vec{n}\cdot\vec{I}_i. \tag{10}$$

Note that the vector

$$\vec{I}_i = I_i(\vec{j}).\vec{i} + I_i(\vec{i}).\vec{j} \tag{11}$$

characterized that the migration of j-th system agents in an arbitrary point $\vec{r}$ of the area G, i.e. it is a vector of migration flow for the system of agents at the point $\vec{r}$.

Let us now consider an arbitrary circle $B_{(\vec{r},\varepsilon)} \subseteq G$. Let the boundary of the circle ∂B is outside normal $\vec{n}$. Then the time dt in this round will occur the $\int_B dN_i^{vol.} = \int_B R_i dSdt$ number of individuals from i-th system agents, and through his border will go $\int_{\partial B} \vec{I}_i(-\vec{n})dldt$ number individuals of the same system agents. Then the total change in the number of individuals (by the system) in the field is

$$\frac{\partial}{\partial t}\int_B dN_i = \int_B \frac{\partial \rho_i}{\partial t} dS. \tag{12}$$

Now, we consider the (16), we obtain the integral equation:

$$\iint_B \frac{\partial \rho_i}{\partial t} dS = -\int_{\partial B} \vec{n} \cdot \vec{I}_i dl + \iint_B R_i dS. \tag{13}$$

When applying the Gauss theorem to the integral along the border in the above we get:

$$\iint_B \left(\frac{\partial \rho_i}{\partial t} + \nabla \cdot \vec{I}_i - R_i \right) dS = 0. \tag{14}$$

Note that the above equation applies to any round of explosives.

Finally, if we consider a continuous integrand, we can shrink the volume in point. Thus we get the equation:

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot \vec{I}_i = R_i, \quad (i = 1,...,n). \tag{15}$$

We see that the parameters of the state of the above system are ρ_i and $\vec{I}_i$. If we have no migration, i.e. $\vec{I}_i = 0$ it follows that $\rho_i = const. \Rightarrow \nabla \rho_i = 0$. The opposite statement is also true. Indeed, if the density is constant, we have no migration, i.e. $\vec{I}_i = \vec{0}$.

This shows that the parameters of the state we can take $\vec{I}_i$ and $\nabla \rho_i$. The simplest (and most natural) case is that in which the ratio between them is non-linear (in analogy to the law of Hooke elastic body), or:

$$\vec{I}_i = -\sum_{j=1}^{n} D_{ij} \nabla \rho_j. \tag{16}$$

Thus the system (15), we get:

$$\frac{\partial \rho_i}{\partial t} - \sum_{j=1}^{n} D_{ij} \nabla \rho_j = R_i, \quad (i = 1,...,n). \tag{17}$$

If we take the function of increasing the number of individuals (in the absence of migration) from the model Dimitrova -Vitanov [10,11,12]:

$$\begin{aligned} R_i = r_i^0 \rho_i [1 - \sum_{j=1}^{n} \left(\alpha_{ij}^0 - r_{ij} \right) \rho_j + \\ - \sum_{j,k=1}^{n} \alpha_{ij}^0 \left(\alpha_{ijk} - r_{ik} \right) \rho_j \rho_k - \sum_{j,k,l=1}^{n} \alpha_{ij}^0 r_{ik} \alpha_{ijl} \rho_j \rho_k \rho_l]. \end{aligned} \tag{18}$$

The above expression is obtained for the following reasons:

We know [4] that the generalized equation of Lotka - Voltera is:

$$\frac{dN_i(t)}{dt} = r_i N_i(t) \left[1 - \sum_{j=1}^{n} \alpha_{ij} N_i(t)_j \right], \quad (i = 1,...n), \tag{19}$$

where is the number of the agents of the i-th agent system, is ratio of increase in the number of individuals i-th system agents, and are the coefficients of the interaction, showing how j-th system agents affects the i-th system.

In this model, the coefficients of the growth factors and interactions are constants. However, the change in the number of agents from a system agent leads to a change in the number of meetings between them. This in turn can lead to a change of both the ratios of increase and rates of reaction. This is a manifestation of the adaptation of a separate system of agents to changes in environmental conditions. Considering this adaptation Dimitrova and Vitanov offer the following dependencies of the coefficients and the number of agents [10]:

$$r_i = r_i^0 \left[1 + \sum_{j=1}^{n} r_{ij} N_j + \sum_{j,k=1}^{n} r_{ijk} N_j N_k + \sum_{j,k,l=1}^{n} r_{ijkl} N_j N_k N_k + \ldots \right], \tag{21}$$

$$\alpha_{ij} = \alpha_{ij}^0 \left[1 + \sum_{k=1}^{n} \alpha_{ijk} N_k + \sum_{k,l=1}^{n} \alpha_{ijkl} N_k N_l + \sum_{k,l,m=1}^{n} \alpha_{ijklm} N_k N_l N_m + \ldots \right]. \tag{22}$$

Dimitrova and Vitanov limited to the simplest, yet most natural form of the above relations, keeping only linear members. Thus, the dependence of the coefficients of the growth factors and the number of the reaction of members of the systems of agents obtain:

$$r_i = r_i^0 \left[1 + \sum_{j=1}^{n} r_{ij} N_j \right], \tag{23}$$

$$\alpha_{ij} = \alpha_{ij}^0 \left[1 + \sum_{k=1}^{n} \alpha_{ijk} N_k \right]. \tag{24}$$

In this case the adaptation is determined only by the number of agents of the system and is connected with both the binary interaction as well as interaction of the higher order.

Taking into account (3) and (4) for the system of ordinary differential equations for modeling the dynamics of systems of agents in the presence of adaptation gets the form:

$$\frac{dN_i}{dt} = r_i^0 N_i [1 - \sum_{j=1}^{n} \left(\alpha_{ij}^0 - r_{ij} \right) N_j + \sum_{j,k=1}^{n} \alpha_{ij}^0 \left(\alpha_{ijk} - r_{ik} \right) N_j N_k - \sum_{j,k,l=1}^{n} \alpha_{ij}^0 r_{ik} \alpha_{ijl} N_j N_k N_l].$$

We see that the resulting model is quite flexible because of the large number of factors in it and because of the possibility of these coefficients depend on the number of agents in the system. Thus, as a special case of the above model are obtained from many references known until. Now, we take the functions model Dimitrova - Vitanov, we get the system of equations:

$$\begin{aligned} &\frac{d\rho_i}{dt} - \sum_{j=1}^{n} D_{ij} \Delta \rho_i = \\ &= r_i^0 \rho_i [1 - \sum_{j=1}^{n} \left(\alpha_{ij}^0 - r_{ij} \right) \rho_j + \sum_{j,k=1}^{n} \alpha_{ij}^0 \left(\alpha_{ijk} - r_{ik} \right) \rho_j \rho_k - \sum_{j,k,l=1}^{n} \alpha_{ij}^0 r_{ik} \alpha_{ijl} \rho_j \rho_k \rho_l]. \end{aligned} \tag{25}$$

Fixed points of the above system equations are the respective stationary densities, i.e. (except density at zero) are the solutions of algebraic equations:

$$1-\sum_{j=1}^{n}\left(\alpha_{ij}^{0}-r_{ij}\right)\rho_j+\sum_{j,k=1}^{n}\alpha_{ij}^{0}\left(\alpha_{ijk}-r_{ik}\right)\rho_j\rho_k-\sum_{j,k,l=1}^{n}\alpha_{ij}^{0}r_{ik}\alpha_{ijl}\rho_j\rho_k\rho_l=0.$$

Now, if we consider deviations Q_i from the stationary densities ρ_i^0 , i.e. if you do a substitution:

$$\rho_i(x,y,t)=Q_i(x,y,t)+\rho_i^0,\ i=1,...,n.$$

We get a system of parabolic equations:

$$\begin{aligned}&\frac{d\rho_i}{dt}-\sum_{j=1}^{n}D_{ij}\Delta\rho_i=r_i^0(Q_i+\rho_i^0)[1-\sum_{j=1}^{n}\left(\alpha_{ij}^{0}-r_{ij}\right)Q_j+\\&+\sum_{j,k=1}^{n}\alpha_{ij}^{0}\left(\alpha_{ijk}-r_{ik}\right)(Q_jQ_k+Q_j\rho_k^0+Q_k\rho_j^0)+\\&+\sum_{j,k,l=1}^{n}\alpha_{ij}^{0}r_{ik}\alpha_{ijl}(Q_jQ_kQ_l+Q_jQ_k\rho_l^0+Q_kQ_l\rho_j^0+\\&+Q_lQ_j\rho_k^0+Q_j\rho_k^0\rho_l^0+Q_k\rho_l^0\rho_j^0+Q_l\rho_j^0\rho_k^0)].\end{aligned}\tag{26}$$

2. The dispersion relation for the case of two human populations

Now let us consider the case of two human populations. Will assume that we have two spatial dimensions that correspond to parts of the ground installation:

$$\begin{aligned}&A_{ij}=r_i^0\rho_i^0[\alpha_{ij}^0-r_{ij}+\sum_{k=1}^{2}\alpha_{ik}^0(\alpha_{ikj}+r_{ij})\rho_k^0+\\&+\sum_{k,l=1}^{2}(\alpha_{ij}^0r_{ik}\alpha_{ijl}+\alpha_{ik}^0r_{ij}\alpha_{ikl}+\alpha_{ik}^0r_{il}\alpha_{ikj})\rho_k^0\rho_l^0],\ i=1,2.\end{aligned}$$

Thus, the model system of equations (26) in the linear case we obtain a system of two equations:

$$\frac{\partial Q}{\partial t}+\sum_{j=1}^{2}D_{ij}\Delta Q_j+\sum_{j=1}^{2}A_{ij}Q_j=0,i=1,2.\tag{27}$$

Differentiating parts during the first equation $(i=1)$ of the above system, and replace it in the second $(i=2)$ get the equation of the fourth order:

$$\begin{aligned}&\frac{\partial^2Q}{\partial t^2}+(D_{11}+D_{22})\Delta\frac{\partial Q}{\partial t}+(D_{11}D_{22}+D_{12}D_{21})\Delta^2Q+\\&+(D_{22}A_{11}+D_{11}A_{22}-D_{12}A_{21}-D_{21}A_{12})\Delta Q+\\&+(A_{11}+A_{22})\frac{\partial Q}{\partial t}+(A_{22}A_{11}-A_{12}A_{21})Q_j=0.\end{aligned}\tag{28}$$

If I do the above conversions and the second equation of (27), we'll get to the equation, i.e. receive the same equation for both populations.

Firstly, we consider a simple case, where:

$$D_{ii} = A_{ii} = 0, \quad i = 1,2, \tag{29}$$

i.e. ignore the influence of the very population on themselves and believe that migration is only due to the interaction between the two populations. Thus we get the equation:

$$\frac{\partial^2 Q}{\partial t^2} - D_{12}D_{21}\Delta^2 Q - (D_{12}A_{21} + D_{21}A_{12})\Delta Q - A_{12}A_{21}Q = 0. \tag{30}$$

If we substitute

$$Q = e^{i(\vec{k}\cdot\vec{r} - \omega t)} \tag{31}$$

in the above equation we get the following dispersion relation:

$$\omega^2 = -D_{12}D_{21}k^4 + (D_{12}A_{21} + D_{21}A_{12})k^2 - A_{12}A_{21}, \tag{32}$$

or

$$\omega = \sqrt{(A_{12} - D_{12}k^2)(D_{21}k^2 - A_{21})}. \tag{33}$$

Then the group velocity of the wave (at a fixed value to k_0) we get:

$$\left(\frac{\partial\omega}{\partial k}\right)_{k_0} = k_0\left(D_{12}\sqrt{\frac{A_{21} - D_{21}k_0^2}{D_{12}k_0^2 - A_{12}}} - D_{21}\sqrt{\frac{D_{12}k_0^2 - A_{12}}{A_{21} - D_{21}k_0^2}}\right). \tag{34}$$

Let us now consider the more general case of equation (28). Let now the frequency is complex, i.e. $\omega = \alpha - i\beta$, where α and β are constants the fly, i.e.:

$$Q = e^{i(\vec{k}\cdot\vec{r} - \omega t)} = e^{i[\vec{k}\cdot\vec{r} - (\alpha - i\beta)t]} = e^{-\beta t}e^{i(\vec{k}\cdot\vec{r} - \alpha)t}. \tag{35}$$

Then by performing the actions described in the previous case we get:

$$\beta = \frac{1}{2}\left[A_{11} + A_{22} - \left(D_{11} + D_{22}\right)k^2\right], \tag{36}$$

$$\alpha^2 = -\left[D_{12}D_{21} + \left(\frac{D_{11} - D_{22}}{2}\right)^2\right]k^4 +$$

$$+\left[\frac{1}{2}\left(A_{11} - A_{22}\right)\left(D_{11} - D_{22}\right) + D_{12}A_{21} + D_{21}D_{12}\right]k^2 + \tag{37}$$

$$-\left(\frac{A_{11} - A_{22}}{2}\right)^2 - A_{12}A_{21}.$$

Finally, we will take special point P with coordinates (assume that $\mathrm{r}_{11}\mathrm{r}_{22} \neq \mathrm{r}_{12}\mathrm{r}_{21}$):

$$\lambda_1 = \frac{\mathrm{r}_{12} - \mathrm{r}_{22}}{\mathrm{r}_{11}\mathrm{r}_{22} - \mathrm{r}_{12}\mathrm{r}_{21}},$$

$$\lambda_2 = \frac{\mathrm{r}_{21} - \mathrm{r}_{11}}{\mathrm{r}_{11}\mathrm{r}_{22} - \mathrm{r}_{12}\mathrm{r}_{21}}. \tag{38}$$

We make:

$$\mathrm{r}_{11} = \mathrm{r}_{22} = \mathrm{r}, \mathrm{r}_{12} = \mathrm{r}_{21} = -s, (r, s > 0).$$

Then we get that:

$\lambda_1 = \lambda_2$, and if $r < s$, then:

$$\lambda = \lambda_i = \frac{1}{s-r} > 0. \tag{39}$$

Now appending and

$\alpha_{ij}^0 = \alpha, \alpha_{ijk} = \beta$.

When for A_{ij} coefficients, we obtain:

$$A_{ij} = \lambda r_i^0[\alpha - r_{ij} + 2\alpha(\beta + r_{ij})\lambda + 4\alpha\beta(r_{i1} + r_{ij} + r_{i2})\lambda^2] \tag{40}$$

Thus, for the dispersion relation in this case we get:

$$\omega = \sqrt{\{r_1^0\lambda[\alpha - s + 2\alpha(\beta + s)\lambda + 4\alpha\beta(2s - r\lambda^2)] - D_{12}k^2\}} \times \\ \times\sqrt{\{D_{21}k^2 - r_2^0\lambda[\alpha - s + 2\alpha(\beta + s) + 4\alpha\beta(2s - r)\lambda^2]\}} \quad . \tag{41}$$

3. Conclusion

In this paper, we discuss a model system of parabolic partial differential equations for description of the spatio-temporal dynamics of interacting human populations (The populations can be any social, ethnic, religious group of people, etc.). It is the presence of spatial dimensions of our model allows us to explore and movements of members of populations in space. From a practical point of view Eq. (6) can be used for eventual prognosis and control of the human immigration, for example in the borders of a specific country. Each country can be separated into a number of regions for which some data on the number of immigrants can be found. Birth and death rates are often available from statistical data. From these sources estimations of the coefficients can be obtained. The most difficult task is the determination of the coefficients and, but they can be obtained by standard or appropriate statistical procedures as well. After determining the coefficients we can simulate Eq. (25) with known or appropriate initial conditions. Thereby we can obtain an approximate picture of the human immigration in the observed region for some period of time.

On the other hand , we can get some information about immigration flows in the regions without numerically solving the system of equations , but based on the qualitative analysis of only partial differential equations. For the purpose here to discuss in detail a simple case of migration of the population in the case of two human populations. In this case , the system of the model is reduced to only one -dimensional nonlinear PDE fourth order , which can be prepared just analytical solution . This equation describes the evolution of the spatial density of the individuals in both populations. Solitary waves can travel through the system , if the amplitude of such density is high. In addition , using an appropriate method , you will get exactly special analytical solution of the model equation . This solution describes the nonlinear solitary waves whims and expression of the distribution of the change in density in space.

Finally, we note that the general form of the model, presented by Eq. (25) can be used for description of dynamics of another population, for example populations of animals, viruses, cells, etc.

Acknowledgment

We would like thank to the "ISCH COST Action IS1104 the EU in the new complex geography of economic systems: models, tools and policy evaluation" for the support of our research.

References

1. Витанов Николай К., З. Димитрова, С. Панчев. Популационна динамика и национална сигурност. 2005: Академично издателство "Марин Дринов" на БАН, 2005.
2. Витанов Николай К., З. Димитрова, С. Панчев. Социална динамика без формули. София: Академично издателство "Марин Дринов" на БАН, 2008.
3. Димитрова З. И., Н. К. Витанов, Популационната динамика и неочакваните и приложения. Наука, 12, 27.
4. Abramson G. Ecological models of extinction. Phys. Rev. E 55, 785
5. Anderson R. M., R. May (1992). Infectious diseases of humans. Dynamics and control. Oxford University Press, Oxford.
6. Arneodo A., P. Coullet, J. Peyraud, C. Tresser. Strange attractors in Volterra equations for soecies competition. J. Math. Biology 14, 153. 1982.
7. Arneodo A., P. Coullet, C. Tresser. Occurrence of strange attractors in three-dimensional Volterra equations. Physics Letters A 259-263, 1982.
8. Brauer, F., C. Castillo-Chavez, Mathematical Models in Population Biology, Springer, Berlin, 2001.
9. Cai C. Q., Y. K. Lin (2004). Stochastic analysis of the Lotka-Volterra models for Ecosystems. Phys. Rev. E, 70, 041910.
10. Dimitrova, Z. I., N. K. Vitanov, Influence of adaptation on the nonlinear dynamics of a system of competing populations. Phys. Lett. A 272, 368-380, 2000.
11. Dimitrova, Z. I., N. K. Vitanov, Dynamical Consequences of Adaptation of Growth Rates in a System of Three Competing Populations. J. Phys. A: Math Gen., 34, pp. 7459-7473, 2001.
12. Dimitrova Z. I., N. K. Vitanov, Adaptation and its impact on the dynamics of a system of there competing population. Physica A, 300, 91-115, 2001.
13. Dimitrova, Z. I., N. K. Vitanov, Chaotic pairwise competition, Theor. Population Biol. 66, pp. 1–12, 2004
14. Dimitrova Z. I., Vitanov N. K. (2005). Shilnikov chaos in a generalized system for
15. modelling dynamics of competing populations. Compt. rend. Acad. bulg. Sci., 58, N 3, 257-264.
16. Dimitrova Z. I., N. K. Vitanov, I.P.Jordanov (2005). On nonlinear dynamiks of agent
17. systems: A case of competing population. Proceedings of the 10th Jubilee
18. National Congress of Theoretical and Applied Mechanics, 41-48.
19. Gardiner, G. W., Handbook of Stochastic Methods for Physics, Chemistry, and the Natural
20. Sciences. Springer, Berlin, 1983.
21. Jordanov I. P. On the nonlinear waves in (2+1)-dimensional population systems. Comp. rend. Acad. Sci. Bulg. 61, 307-314, 2008.
22. Jordanov I. P. Nonlinear waves caused by diffusion of population members. Comp. rend. Acad. Sci. Bulg. 62, 33-40, 2009.
23. Kaplan, D., L. Glass, Understanding Nonlinear Dynamics, Springer, New York, 1995.
24. Kiel, L. D. and E. Elliott (eds.), Chaos Theory in the Social Sciences: Foundations and Applications, The University of Michigan Press, Ann Arbor, 1996.
25. Malfliet, W., Solitary Wave Solutions of Nonlinear Wave Equations. American Journal of Physics, 60, pp. 650 – 654, 1992.
26. Martinov, N., N. Vitanov, On the Self-consistent Thermal Equilibrium Structures in Two-dimensional Negative-temperature Systems. Canadian Journal of Physics, 72, pp. 618–624, 1994.
27. Martinov, N., N. Vitanov, New Class of Running-wave Solutions of the 2+1-Dimensional Sine - Gordon Equation. J. Phys. A: Math. Gen., 27, pp. 4611 – 4618, 1994.

28. Montroll, E. W., W. W. Badger, Introduction to Quantitative Aspects of Social Phenomena, Gordon and Breach, New York, 1974.
29. Panchev S., N. K. Vitanov (2008). Mathematical models of intergroup conflicts. Compt. rend. Acad. sci. Bulg. 61, 993-1002.
30. Puu T. (2004). Nonlinear Economic Dynamics. Lec. Notes in Econ. and Math Systems. 336, Springer - Verlag.
31. Richmond P., L. Sabatelli (2004). Peer pressure and generalized Lotka-Volterra models. Physica A, 344.
32. Richmond P., L. Sabatelli (2004). Langevin processes, agent models and socio-economic systems. Physica A,336, 27.
33. Roughgarden J. (1975). A Simple model for population dynamics in stochastic environments. American Naturalist, 109, 713-713.
34. van Kampen, N. G., Stochastic Processes in Physics and Chemistry, Elsevier, Amsterdam, 2007.
35. Vitanov, N. K., I. P. Jordanov, Z. I. Dimitrova, On nonlinear dynamics of interacting populations: Coupled kink waves in a system of two populations, Commun.Nonlinear Sci. Numer. Simul. 14, pp. 2379–2388, 2009.

Predictability of Dynamical Systems

Kostadin Sheyretski

UNWE, 1700 Sofia, Bulgaria
sheiretsky@abv.bg

Abstract: The paper discusses linear and nonlinear dynamic systems; analyzes the reversibility and predictability of nonlinear systems in the context of the paradox of time and the Poincaré-Zermelo theorem. The possible attractors in dissipative systems and the phenomenon of mixing in phase space are being analyzed.

Keywords: dynamical systems, chaos theory.

Deterministic systems

One process is called deterministic if all his future course, and his entire past, is uniquely determined by the present moment. The set of all possible states of the process is called phase space. Deterministic systems will be called dynamic systems as well.

Any process or phenomenon in nature, economics, society, etc. can be modeled by a system of differential equations. This model determines the law of evolution of the system. Basically it is a dynamic system which reveals its development in the future by solving the corresponding equations.

An important concept in the theory of dynamical systems is the phase space: many elements of which consist of a set of positions and velocities of the system [2].

The success of linear models

The success of linear models provided a basis for the claim that the linear approach to describe the world around us is sufficient to analyze all the important events of the phenomena. Complicated processes are broken down into a series of simpler, and they in turn can be described by known and well developed analytical and numerical methods. Confirmation of this hypothesis appears to be created by Maxwell's electrodynamics. A striking example is the Schrödinger equation in quantum mechanics [10]. Quantum mechanics incorporates the uncertainty principle of Heisenberg, who claimed that the position and velocity of a particle cannot be both accurately measured. In order to find out where the particle is, it has to be Lighted, but in case of Planck that Einstein proved, cannot use an infinitesimal amount of light because it moves in portions - quanta. This quantum will pass a certain speed to the particle, and the more accurately we try to determine the position – the greater the speed passed will be. In 1924 Louis de Broglie hypothesized that in the micro world of the particles is possible to attribute both corpuscular and wave properties. Using the analogy between classical mechanics and optics, Erwin Schrödinger's equation brings down the change of the density of probability of finding a particle in space and time.

Despite that success the linear approach, although based on the well-known mathematical constructions, proved inapplicable to study the diversity of the world [5], [8].

Reversibility and irreversibility of the dynamic processes

The success of classical mechanics led to the belief that it can adequately describe any process observed in the universe. Laplace expresses the idea that if being able to determine the initial conditions of all objects in the universe, the further evolution in time of the entire system can be predicted. I have to say that from the present perspective the existence of "Laplace's demon" is not acceptable as the main reasons for this are thermodynamics, quantum mechanics and chaos theory.

In thermodynamic systems it is impossible to describe each motion of the gas component in each particle. Ludwig Boltzmann offers a statistical approach. It establishes a relationship between the degree of disorder in a system and the magnitude of entropy. According to the second law of thermodynamics, entropy in an isolated (one that doesn't exchange energy with the environment) thermodynamic system does not decrease with time, which means that the disorder is growing. Second law of thermodynamics is closely linked to the concept irreversibility.

Belgian physicist Ilya Prigogine defines irreversibility as a creator and an expression of time [7]. In classical physics there is no difference between past and future, there is symmetry between the directions of time – the time paradox. Prigogine contrasts the terms "irreversible" and "reversible", and irreversibility is anchored in time - it cannot be reversed and past cannot happen again present. But as the Hawking says that entropy increases with time because we define the direction of time (arrow of time) is that in which entropy increases [10]. This means that the second law of thermodynamics is a tautology.

German physicist Clausius applies the second law of thermodynamics to the entire universe, considering it as an isolated thermodynamic system. So he concludes that in time the temperature everywhere will align all processes in nature will end - state of complete disorder known as heat death.

In contrast to the irreversibility appears to be a strictly proven theorem and it is also applicable to thermodynamic systems.

Conservative system in place is the Poincaré-Zermelo theorem: Almost anywhere from any region of phase space, moving along the trajectory will return to the same area.

Poincaré-Zermelo theorem is an expression of microscopic reversibility of movement of a system of material points. This is due to the reversibility of the invariance equations of motion with respect to the temporal inversion, i.e. the replacement of t with –t. Experience shows, however, that a closed thermodynamic system, after a sufficiently large time interval passes at steady state, then this state practically does not change with time. Poincaré-Zermelo theorem shows that even very small probability of spontaneous passage of one present in equilibrium system in a condition that is notably imbalances, is distinguishable from zero. The minimum period of time after the expiry of which a practical micro-condition is repeated is called Poincaré cycle. Poincaré cycles are huge and are much larger than the age of the universe [3].

Unpredictability of deterministic processes

Could the deterministic process still to be unpredictable? It turns out that it is possible and it is a characteristic of some of the non-linear systems.

Henri Poincaré draws attention to the possibility of the existence of processes in which small differences in initial conditions can lead to huge changes in the behavior of the system over time.

Very bright example of such a phenomenon is described by the writer Ray Bradbury in the story "A Sound of Thunder."

The main character of the story is one of the organizers of election campaign. After the victory of its candidate, he participated in a Tyrannosaurus Rex safari in the Mesozoic era, as previously is known that the animals that will be killed are going to die in few hours. Hunters return 60 million years back in time. In order not to violate causality, hunters move in certain paths. The hero of the story does not fulfill this condition and crushes without asking a golden butterfly. On his return in 2055 he found that the reality is radically altered, amended the composition of the atmosphere, the spelling of English is different, and has changed people's behavior, and even the political system, and the outcome of that election campaign was involved in character .

Almost simultaneously with the creation of the story, an American scientist Edward Lorenz literally see the image of the butterfly in the image of the phase portrait of a system that explores.

In 1963 the scientist creates a model for predicting the processes in the atmosphere. This is what tells the discovery itself Lorenz: "At one point I decided to repeat some of the calculations in order to clarify in more detail what happens. I stopped the computer, I implemented a series of numbers, which it shortly before produced, and started it back to work. I went out into the hall to drink a cup of coffee and came back after about an hour, during which time the computer was simulated meteorological forecast for about two months. Printed numbers had nothing to do with the previous ones. [...] Blame for what happened is drifted initial rounding errors: they continuously grew and began to dominate in the decision. Speaking of today's language, it was chaos. "[4]

"Predictability: Can the flapping of a butterfly in Brazil can cause a tornado in Texas?" is the title of a lecture by Edward Lorenz, delivered on 29 December 1972 in Washington on 139th conference of the American Association for the Advancement of Science. Initially the title of the lecture used seagull instead butterfly but the chairman of the meeting changed it and so the "Lorenz butterfly" appears and "the butterfly effect" becomes one of the most metaphors of chaos theory.

A major blow to the Laplacian determinism applied Heisenberg, who formulated the uncertainty principle in quantum mechanics. Let me recall that the principle asserts that in nature there is an objective state of a particle with specific values of the coordinates and momentum. This principle explains why processes arising in the nucleus of atoms are generally unpredictable, no matter how much information we have collected.

For the larger scale structures uncertainty principle is not applicable. If the system consists of a large number of degrees of freedom (eg gas in a container), the situation is clear: the system can be described by the equation of motion of each particle, but this task becomes unreal. No modern computer would not solve the problem even for one mole of gas.

What about systems with fewer degrees of freedom? It turns out that while a system is deterministic, i.e. after setting the initial conditions and the behavior is completely determined at any time, it can become unpredictable chaotic. For example, in the game of billiards. We assume that the player fully control your shot and wants to predict the trajectory of the ball that hit after a series of clashes on the billiards table and the other balls. It turns out that even ignore the gravitational attraction of an electron from the edge of the galaxy, the forecast will be wrong now even a minute. The reason lies in the instability of all or almost all the trajectories. [9]

Dissipative systems. Regular and strange attractors of dynamical systems

Dissipative systems are dynamic systems associated with loss of energy. Movements of dissipative systems can be divided into two classes: transient, non-stationary matching process of relaxation from the initial to the marginal variety of states and established , fixed movements, the phase trajectories are fully belong to the limit set. Gravity limit sets are called attractors. Over time from any initial state, situated in an area to relax attractors [6].

Attractors can be points and limit cycles (closed curves) in two dimensional space, but can be and manure in three-dimensional space. It is wound trajectory, the system periodically move. All these attractors are regular.

In the phase space with a dimension of not less than 3 are possible, and another type of attractors. Ruel and Takens call these strange attractors. Dynamic system stuck in zone strange attractor has established aperiodic fluctuations and not predictable in time, because it is characterized by sustainable mode of operation. Region of phase space in which it is located is called the attractor region or basin of attraction. An example of such attractor is the Lorenz attractor. Such areas are characterized by exponential divergence unless the orbit and fractional dimension of a set.

Interfusion

In dissipative system, the dimension of which is 3 or more, it is theoretically possible mode of complex non-periodic pulses. This movement is deterministic and is characterized by a vulnerability.

Let in the phase space of dissipative systems take small sphere of points - initial conditions and examine what will happen over time. Because the system is dissipative, if there is a resistance of the trajectories, the field will shrink permanently, in the limit $t \to \infty$ where the radius decreases to zero.

In dynamic mode, if the system is linear and conservative, the phase volume will increase indefinitely. In a system which is nonlinear and dissipative, instability leads to trajectories “running away" from one another, and dissipativity - to a phase volume contraction. The phase displacement in one direction will be stretching, and in the other - shrinking. The degree of contraction should prevail over the degree of stretching, because the system is dissipative. So the considered small sphere interfusions the entire area occupied by the trajectory as a kind of volume and "smears" on it. [1]

Conclusion

Briefly was presented the main paradigms in the study of dynamical systems. Although deterministic, they often find themselves with an unexpected and unpredictable complex behavior, although relatively simple differential equations that define them. Because of the universal nature of the models, the study of nonlinear effects in the evolution of deterministic systems is the key to a better understanding of the world around us, and it sets new philosophical questions, the answer to most of which today is still not found.

References

1. Anishtenko, V. Introduction to Non-linear Dynamics, Moscow, IKI 2002. (in Russian)
2. Arnold, V, I. Ordinary Differential Equations, Ijevsk "Regular and Chaotic Dynamics", 2000. (in Russian)
3. Ivanov, K, Dakova, D. Statistical Physics and Thermodynamics, Plovdiv, Plovdiv's University Publishing House 1995. (in Bulgarian)
4. Lorentz, E. The Essence of Chaos, Sofia, University Publishing House St. Kliment Ohridski, 2000. (in Bulgarian)
5. Malinetskii, G. Mathematical Basis of Synergetics, Moscow, URSS, 2009. (in Rissian)
6. Pantchev, St. A theory of Chaos, Sofia, Academic Publishing House Prof. Marin Drinov, 1996. (in Bulgarian)
7. Prigogine, I. Stengers, I. Order out of Chaos: Man's new dialogue with nature, Moscow, URSS, 2003. (in Russian)
8. Trubetzkov, D. I . Introduction to Synergetics: Fluctuations and waves, URSS 2011. (in Russian)
9. Trubetzkov, D. I. Introduction to Synergetics: Chaos and structure. Moscow, URSS, 2012. (in Russian)
10. Hoking, S. Black Holes and Baby Universes and Other Essays. Sofia, Science and Art 1994. (in Bulgarian).

Time Series Analysis of Successful Economics Regulation Politics

Nikolay K. Vitanov[1], Ivan P. Jordanov[2], Ivan S. Ivanov[3]

[1] Institute of Mechanics, Bulgarian Academy of Sciences, Sofia, Bulgaria
[2,3] Faculty of Applied Informatics and Statistics,Department of Mathematics, UNWE, Sofia, Bulgaria
[2] i_jordanov@email.bg, [3]ivanov_i_s@abv.bg

Abstract: Many social problems can be studied of the methods of nonlinear dynamics, time series analysis and economic statistics. In this article we explore some experimental time series related to the cost of agricultural production. These time series are short and nonstationary so must be analyzed with methods different from those applied to stationary time series. As they are nonstationary we cannot apply the classical methods based on the reconstruction of the phase space of the system. Instead of this we shall use the singular values decomposition (SVD) method. We analyze time series of prices and agricultural production before and after the second oil crisis of 1974 and the second illustrates the effects of government regulation on them. We show that after proper fiscal intervention by the government in agricultural production can be corrected the operations of market mechanisms, despite the sharp change in the prices, changes in output are not as sharp. We find that the increase in the dimensionality of the phase space in the prices reduces the cyclical behavior of the time series. In the time series behavior of the production is back on the time series of prices.

Keywords: nonstationary time series, nonlinear dynamics, singular values decomposition (SVD) method

1. Introduction

In the last few decades is observed deepening economic and demographic problems both in development and in developing countries. Consequently, the world is increasingly going into the deepest economic and social crisis, known until now. On the other hand , a purely market mechanisms already conducted more than two centuries , no longer able to better manage and regulate processes in the economy. Today more than ever it is evident that the so-called liberal model based on the vision of the "invisible hand of the market", we not only bring our entry into economic and demographic crisis, but also deepening. Of course a number of leading economists showed at the end of the last century, the devastating impact of extreme liberalism not only developing but also the world's leading economies [4].

Currently, more and more leading scientists, including economists and sociologists have concluded that in order to ensure not only sustainable, but also gradually developing the processes in society reinforces the role of government regulation. In recent years there has been intensive and the development of mathematical models and methods, as well as adapting the methods of applied mathematics, physics and biology to complex systems such as social, economic, technical systems [2,5,12,13,14]. Today the study of the behavior of such complex systems has become increasingly necessary to be able both to predict their behavior and to influence them.

2. Obtaining Time Series

The vast majority of the time series obtained from experimental (real) economic systems. These time series in most cases are not sufficiently long, and often are not stationary. In these time series for now we can only use the method of calculating the autocorrelation, power spectra and histograms [1, 2, 6, 7, 8, and 9].

An analysis of economic time series, ie time series of real economic system is a standard case - the analyst receives the time series as written and not been consulted in advance what length and density should have this time series. This strongly influences the methods of analysis.

It is important to establish whether the investigated time series are stationary or non-stationary. We have been given time series of real economic system, which are non-stationary and ram it with a short length and density. Therefore we cannot use the methods of reconstruction of the phase space of the system. However, we can extract information about the time series, by SVD (singular values decomposition) method [3, 7, 8, 10, and 11].

The results of the analysis of time series of prices and production of piglets in Japan before and after the second world oil crisis of 1974 clearly demonstrated the correctness of the policy of the Japanese government to regulate the market of piglets in Japan.

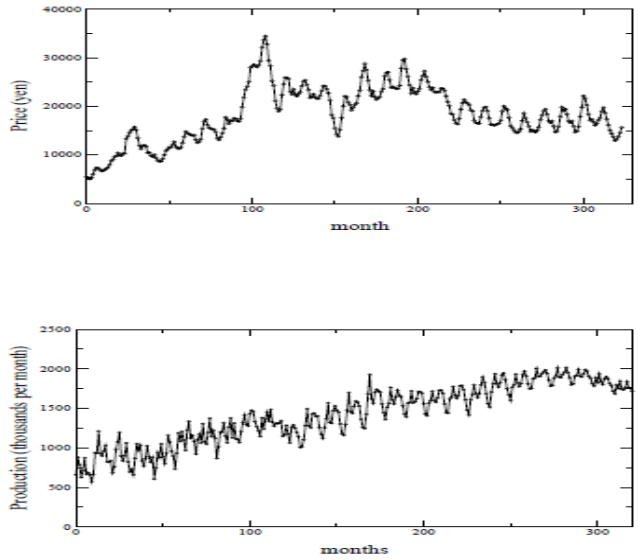


Fig.1 Prices and production of piglets in Japan for the period January 1965 - January 1992

Aiming to support national agricultural production and to protect it from weather and other fluctuations (such as those in oil prices) any government intervene more or less to limit the fluctuations in prices. Opponents of these interventions argue that the market automatically optimizes production and the impact on market mechanisms lead to more harm than good. This argument did not stop the Japanese government from massive law and financial intervention in the agricultural sector after the second oil crisis of 1974.

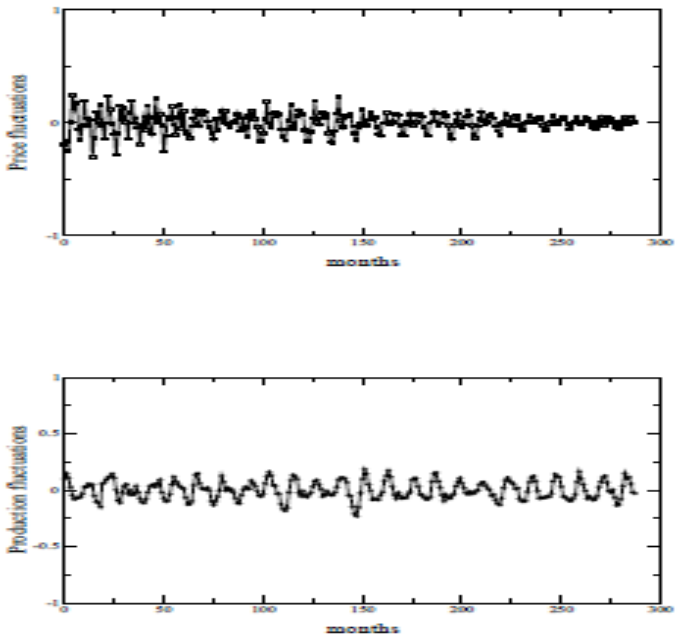


Fig. 2 Fluctuation of prices and production of piglets per month in Japan

Steps have been undertaken in two directions: the introduction of floating recommended purchase prices and floating production subsidies to farmers. The term floating I mean the fact that the value is changed by a decision of the relevant governmental authority in case the market situation. For example, overproduction leads to a reduction of subsidies and insufficient production - to increase the recommended purchase price. As far as the available research time series extend well beyond July 1974 We can see the effects of this government intervention on prices of piglets in Japan.

3. Time Series Analysis

Interestingly, the as can be seen from the figure for the fluctuations, after the second oil crisis of 1974 and the subsequent intervention by the Japanese government to regulate the prices of agricultural output is seen quite a substantial reduction of the fluctuations at the end of the period, while the fluctuations in output is not reduced substantially.

As a preliminary remark, it can be said that the steps the Japanese government were more fiscal, as we have a sharp change in the autocorrelation function of prices after the second oil crisis, changes in output are not as sharp. However, output and noticed changes - Increase in annual cycle at the expense of the semi.

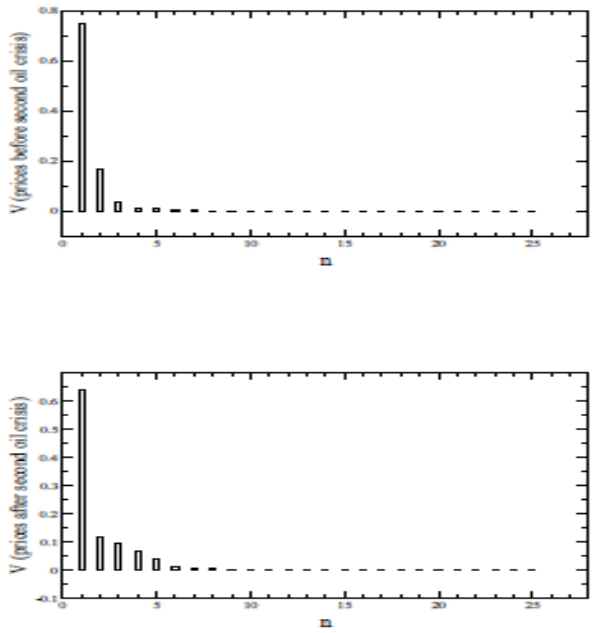


Fig. 3 Part of the total variance of the time series of prices of piglets in Japan before and after the second oil crisis contained in each of the 25 principal components.

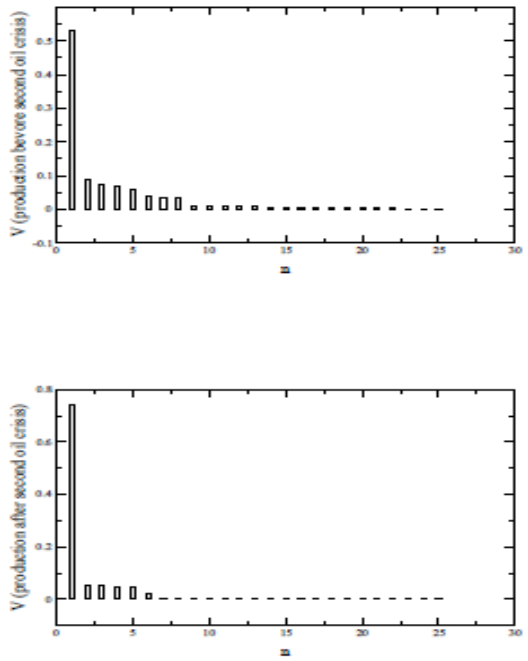


Fig.4 Part of the total variance of the time series for the production of piglets in Japan before and after the second oil crisis contained in each of the 25 principal components.

As to the portion of the total variance of the time series contained in each of the principal components , it can be noted the following: -dominated part of the variance before the crisis was contained in the first three principal components until after the crisis the majority of the variance is already distributed between the first five principal components. In other words, in terms of variance, the number of significant principal components has increased from 3 to 5 as a result of regulatory intervention by the Japanese government in the market for agricultural products. So the dimensionality of the phase space of price increases and dilutes the effect of cyclic behavior is dominated by the second oil crisis, i.e. government intervention leads to a reduction of the intensity of the porcine cycle.

The reverse process takes place in the investigation of the variation in the principal components of the time series of the output. It is seen that after 1974 there is a concentration of variance in the first principal component at the expense of others who have taken a relatively larger part of the variance of the time series before the second oil crisis. This suggests reducing the dimensionality of the phase space that is essential to the dynamics of the time series for the production of piglets.

Idea of the upper limit of the dimensionality of the attractor associated with a time series can be obtained on the nature of the spectrum of singular values. Usually separation was observed in the size of singular values of the two groups. The eigenvectors corresponding to the largest singular values form a space in which the attractor is associated with the time series, while the vectors of smaller singular values are proportional to the noise and fluctuations in the system. What we see in the price of piglets is increasing the number of significant principal components from 3 before the crisis to 5 after the crisis. While the small number of significant principal components shows that the essential dynamics of the system is low dimensional and can be described by a small number of equations.

4. Conclusion

We analyzed the time series of prices and production of the piglets in Japan before and after the second oil crisis of 1974 and for the second we indicating the impact of government regulation. After proper fiscal intervention by the government in agricultural production can be adjusted operation of market mechanisms, despite the dramatic change in prices, changes in output are not as sharp. For the variance of the time series of prices before and after the crisis, i.e. before and after the regulatory intervention is received increased from 3 to 5 of the relevant principal components of the system.

We find that the increase in the dimensionality of the phase space of prices reduces the cyclical behavior in the time series.

After 1974 there is a concentration of variance in the first principal component at the expense of others, which before the second oil crisis have taken a relatively larger part of the variance of the time series.

Acknowledgment

We thank to the Nonlinear time series analysis group from the Max-Planck Institute for the Physics of Complex Systems in Dresden, Germany, for the possibility of using of the TISEAN software package.

We would like to express a special thanks to Professor Sakai of Tokyo University of Agricultural Technologies (TUAT) for kindly provided time series analysis of the evolution of the prices of piglets in Japan.

References

1. Abarbanel H. D. I. Analysis of Haotic Time Series. Springer, New York, 1996.
2. Abarbanel H. D. I., R. Brown, J. J. Sidorowich, L. Sh. Tsirming. The analysys of observed chaotic data in physical systems. Rev. Mod. Phys. 65, 1331-1392, 1993.
3. Albano A. M., J. Muench, C. Schwant, A. I. Mees, P. E. Rapp (1988). Singular-value decomposition and the Grassberger-Procaccia algorithm., Phys. Rev A, 38, 3017.
4. Allais M. (1999). La crise mondial d'aujourd'hui. Edition Clement Juglar, Paris.
5. Ashenfelter K. T., S. M. Boker, J. D. Waddel, N. Vitanov. (2009) Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation. Journal of Experimental Psychology: Human Perception and Performance, 35 1072-1091
6. Blockwell P. A., R. A. Davis. (1995). Time series: theory and methods. Springer, New York.
7. Diks C. (1999). Nonlinear time series analysis: Methods and application. World Scientific, River Edge, New Yersey.
8. Grassberger P., T. Schreiber, C. Shaffrath. (1991) Nonlinear time sequences analysis. Int. J. Bif. Chaos, 1, 521-547.
9. Hamilton J. D. (1989). A new approach to the economic analysis of nonstatitonary time series and the business cycle. Econometrica ,57, 357-384.
10. Hegger R., H. Kantz, Improved false nearest neigbours method to detected determinism in time series data. Phys. Rev. A 60, 4970, 1999.
11. Hegger R., H. Kantz, T. Schreiber. Practical implementation of nonlinear time series methods: The TISEAN packade. CHAOS 9, 413, 1999.
12. Sakai K., S. Managi, N. K. Vitanov, K. Demura (2007). Transition of chaotic motion to linit cycle by intervention of economic policy: An empirical analysis. Nonlinear Dynamics, Psyhology and Life Sciences. 11, 253 - 265.
13. Vitanov N. K., K. Sakai, I. P. Jordanov, S. Managi, K. Demura (2007). Analysis of a Japan government intervention on the domestic agriculture market. Physica A, 382, 330 - 335.
14. Vitanov N. K., K. Sakai, E. D. Yankulova (2005). On correlations and fractal characteristics of time series. J. Theor. Appl. Mech., 35, 2, 73-90.

Using of Reversible Structural Models for Modelling Objects with Distributed Parameters

Volodymyr Fedorchuk[1], Alexander Mahovych [2]

[1] K-PNU, Ivan Ohienko Str., 32300 Kamianets-Podilskyi, Ukraine
fedvlad@ukr.net

[2] K-PNU, Ivan Ohienko Str., 32300 Kamianets-Podilskyi, Ukraine
umismag@gmail.com

Abstract. A structural reversible model of the distributed nonlinear links of the drilling unit is obtained. Proposed model takes into account: drill sting inhomogeneity caused by using drilling pipes of different types; derrick load-caused deformation and its dynamics; resistance forces, caused by interaction of drilling liquid with the drill string's walls; popping force; drilling liquid column flywheel action; resistance forces, caused by the chisel interaction with formation. Demonstrated modeling results indicate that the derived model can be successfully used: on the drilling device composition calculation stage; in control system for maintaining optimal behaviors by drilling; for round-trips operations optimization.

Keywords. Structural reversible model, drilling unit, distributed links.

1. Introduction

In shaft construction there is a necessity to maintain high technical and economic indexes of a drilling process. Especially crucial this problem gets in construction of penetrating and superdeep holes. As it is known, with depth hardness of rocks and temperature in the stope zone, power consumption of rock breaking considerably increases, flush fluid input is handicapped. All it degrades rock breaking process, slashes durability of a drilling rock-destruction tool and other links of the drilling unit.

Increasing drilling technical and economic indexes can be achieved by improving of rock breaking process dynamics by transmitting of mechanical efforts through a drilling string from a surface to a stope and backwards, which, in turn, requires using computer models of drilling unit links in the automated control system [1]. In particular it concerns links with the distributed arguments which are represented in the form of the distributed links. In the drilling unit a link with the most expressed properties of allocation of arguments is the drilling string.

Application of a drilling string computer model in drilling process automated management enables to solve some of the relevant problems: supply of an optimum bit loads by taking into account drill column dynamic components' interaction with a shaft stope; reduction of time of elevating operations by means of admissible load arguments automatic support with the interaction of a string with flush fluid taken into account; reduction of energy consumptions as a result of efficient environment control of ground-level system.

The modern drilling unit is a difficult complex of the miscellaneous equipment consisting of motor engines, equipped with control systems; mechanisms for tool lifting, descending and rotating; pumps for a flush fluid priming; machineries for upward-downward

operations mechanization and automation; a derrick for holding of the tackle dodge and placement of drill-pipes [2]. During shaft penetration in a drill column there are pressure waves which transfer energy of oscillatings and determine dynamics of process of rock breaking. Insufficient learning of the processes which are appearing during drilling, neglecting of dynamic components at control lead to control deterioration, and on occasion to appearing of self-oscillations and, as consequence, to contingency situations.

It is necessary to notice that the analysis of the processes, which are originating at shaft drilling, is possible only under condition of a comprehensive approach when the system is viewed with taking into account interaction of motor engines, gears, executive elements and carrying structures. Thus, it is necessary to deal with models of elements of the different physical nature which have essential differences in their mathematical specifications [3, 4]. The analysis of such systems by means of traditional methods when the set of equations containing ordinary differential equations, the equations in private derivatives, the nonlinear differential equations, the integrated and integro-differential equations dares, is handicapped, first of all, by absence of methods of their mutual solution.

The approach is offered when difficult non-uniform dynamic system is decomposed into more simple structural members that gives the chance to reduce a complex problem solution into solution of certain quantity of simple problems. Application of a principle of decomposition allows introducing complex model in the form of certain pattern consisting of simpler blocks.

At numerical realization for each structural member of macromodel a possibility is proposed to choose the most effective method which determines algorithm used in the applicable programmatic modulus. It is possible to call such approach structurally-algorithmic as to a certain structural member the algorithm is put in conformity. In the end result application of the structural approach gives the chance to synthesize models of difficult systems from a certain baseline set of computer module.

At a stage of models approbation as modelling environment it is expedient to use interactive toolkit for modelling, cloning and analysis of dynamic systems — Simulink/MatLAB. As in serial exploratory programs of computer mathematics, and in particular in Simulink/MatLAB, there are no means for modelling of links with the distributed arguments, and also links with dynamic nonlinear dependences, there is a problem of development on the basis of existing libraries of macromodels of the sample distributed and nonlinear dynamic links of the drilling unit.

Members with the distributed arguments cause significant difficulties at numerical realization in drilling unit blocks modelling. First of all it concerns drilling strings which are non-uniform as they consist of the miscellaneous equipment (steel and weighted drill-pipes, centralizers, bushings, calibrators, expanders, catchers of a slime, etc.).

Also an element with the distributed arguments is the pole of flush fluid which moves through a drilling string to a stope. Necessity of taking into account dynamics of move of flush fluid for a drilling string is caused also by the fact that flush fluid at a slide drilling is used for power transmission to the motor engine, and control of the rotational moment on a bit is carried out through hydraulic connection.

It is possible to view the drilling string, taking into account a number of assumptions, as an elastic rod with the distributed mass, resiliency and a viscous friction. Axled reaction of a stope and bit is applied to a string downhole end, and lengthwise — the distributed forces of weight, a viscous friction, flywheel action. At a drill column there are three kinds of strain — torsions, a flexure and an extension-compression. For obtaining of high-grade results of modelling it is necessary to allow for interaction of formation, a cone rock bit, a downhole

motor and a drilling string. define corporate sustainability is not an easy task. Even till now terms like CSR (Corporate social responsibility) and corporate citizenship are continue widely to be used. Nowadays, however, all these concepts are replaced by the broader term corporate sustainability. In the beginning it was not customary to define sustainability at corporate level.

The purpose of this report is to review the most common concepts of corporate sustainability in order to outline basic qualitative aspects of the most popular views. The ultimate goal is to find promising directions for building quantitative models of corporate sustainability at a later stage.

The most significant consideration is that these quantitative models had to be built on the basis of information available to the organization. This is information that is collected on a daily basis in corporate databases as a result of current activities.

2. Constructing of Mathematical Model

We will consider ways of the mathematical specification of drilling string longitudinal and rotary movement dynamics.

Velocity of offset of sections of a string $V(x, t)$, caused by operating of linear exterior force $F(x, t)$ is described by the differential equation in private derivatives

$$m(x)\frac{\partial^2 V(x,t)}{\partial t^2}+h(x)\frac{\partial V(x,t)}{\partial t}-k\frac{\partial^2 V(x,t)}{\partial x^2}=\frac{\partial F(x,t)}{\partial t}-\frac{\partial T(x,t)}{\partial t} \quad (1)$$

with boundary and initial conditions accordingly (2) and (3)

$$\frac{\partial V(l,t)}{\partial x}=0;\ \frac{\partial V(0,t)}{\partial x}=0, \quad (2)$$

$$\frac{\partial V(x,0)}{\partial t}=0;\ V(x,0)=0, \quad (3)$$

where $m(x)$ — linear mass of string; $h(x)$ — linear coefficient of resistance of a string; k — a resilience factor; $F(x, t)$ — linear exterior force; $T(x, t)$ — linear force of gravity; $V(x, t)$ — travelling speed of section of a string.

For the specification statement of dynamics of rotary movement of a string the equation is used

$$J\frac{\partial^2 \omega(x,t)}{\partial t^2}+\tau\frac{\partial \omega(x,t)}{\partial t}-GJ_p\frac{\partial^2 \omega(x,t)}{\partial x^2}=\frac{\partial M(x,t)}{\partial t} \quad (4)$$

with boundary and initial conditions accordingly (5) and (6)

$$\frac{\partial \omega(l,t)}{\partial x}=0;\ \frac{\partial \omega(0,t)}{\partial x}=0, \quad (5)$$

$$\frac{\partial \omega(x,0)}{\partial t}=0;\ \omega(x,0)=0, \quad (6)$$

where ω — rotating speed of a string; J — inertia moment; G — shift modulus; M — the torsion moment; τ — coefficient of resistance of rotary movement; J_p — polar inertia moment.

However, if a drilling string is non-uniform, and also at nonlinear posing of a problem, approach application when the source system of the differential equations is solved in private derivatives, is inefficient.

Use of structural models in the form of multimass systems with the concentrated masses is offered [6]. Having applied a method of straight lines to the equations (1) and (4), we will receive two systems from n differential equations each

$$m_i \frac{\partial^2 V_i}{\partial t^2} + h_i \frac{\partial V_i}{\partial t} - k_i (V_{i-1} - 2V_i + V_{i+1}) = \frac{\partial F_i}{\partial t} - \frac{\partial T_i}{\partial t}, \qquad i = 1, 2, \ldots n, \qquad (7)$$

$$J_i \frac{\partial^2 \omega_i}{\partial t^2} + \tau_i \frac{\partial \omega_i}{\partial t} - GJ_{pi} (\omega_{i-1} - 2\omega_i + \omega_{i+1}) = M_i, \qquad i = 1, 2, \ldots n, \qquad (8)$$

where V_i — travelling speed of the i-th member; F_i — the exterior force operating on the i-th member; T_i — the force of gravity operating on the i-th member; h_i — coefficient of resistance for the i-th member; k_i — a resilience factor for the i-th member; ω_i — rotating speed of the i-th member of a string; M_i — the moment of torsion operating on the i-th member τ_i — coefficient of resistance to rotary movement of the i-th member of a string; J_{pi} — polar inertia moment of the i-th member of a string.

After integration of set of equations (7) and (8) we will receive

$$m_i \frac{\partial^2 x_i}{\partial t^2} + h_i \frac{\partial x_i}{\partial t} - k(x_{i-1} - 2x_i + x_{i+1}) = F_i - T, \qquad i = 1, 2, \ldots n, \qquad (9)$$

$$J_i \frac{\partial^2 \varphi_i}{\partial t^2} + \tau_i \frac{\partial \varphi_i}{\partial t} - GJ_{pi} (\varphi_{i-1} - 2\varphi_i + \varphi_{i+1}) = M_i, \qquad i = 1, 2, \ldots n, \qquad (10)$$

where x_i — coordinate of the i-th member; φ_i — an angle of displacement of the i-th member.

The received model in the form of system of the differential equations allows to make decomposition of source model on variables x and φ on n links. If to hold decomposition further at level of the received links it is expedient to dedicate in its pattern of magnitude which one have physical sense. It will give the chance to realize dynamic nonlinear dependences through static (for example, use of nonlinear dependence in the form of hysteresis loop allows to recreate in structural model the effect of loss of a part of energy at a plastic deformation owing to internal friction).

After equivalent transformings of systems (9) and (10) for realization of the i differential equations we will receive a structural link in the form of a subsystem of the simulink-model shown on fig. 1.

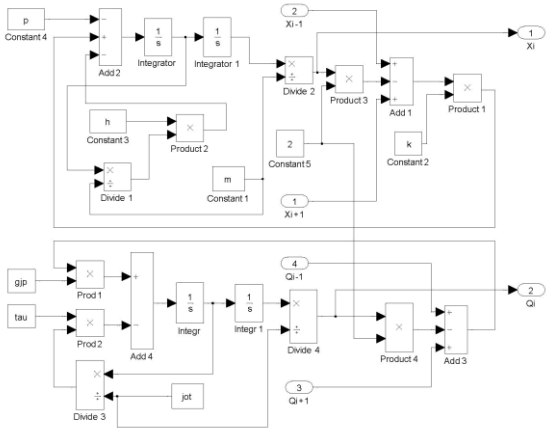


Fig. 1. Structural realization of a link for modelling of a lease of a string with the concentrated mass at forward and rotary movement

The received subsystem compounds a ground for synthesis of the structural computer model realizing set of equations (9) — (10). Joint of subsystems is carried out according to connections through arguments in systems of the differential equations (fig. 2).

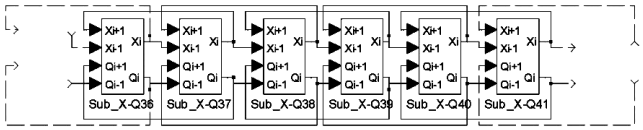


Fig. 2. Structural realization of a lease of a string at approximating in the concentrated masses

The structural model of a drill column synthesized this way has a number of the relevant properties. First, string digitization on links with the concentrated masses gives the chance to hold for each link specification of model by the account of certain features on string leases, including nonlinear dependences (for example, variation of temperature with depth growth, and also other arguments of model, temperature-dependent). Secondly, at the expense of the organization of straight lines and feedback between links, the model has property of convertibility that gives the chance to carry out input impacts and to receive interesting results at level of any structural link of model.

At constructing of a drill column model we will make the assumption that it is homogeneous on final quantity of leases of nonzero length. Then within one such lease we will have invariable arguments. We will examine some of them.

Inertia moment of a lease of a string of linear mass *m* and length *l* is found from expression

$$J = 0,5ml(R^2 + r^2),$$

where *R* and *r* are, accordingly, exterior and internal radiuses of a string.

Polar inertia moment of a lease of a string is calculated by the formula

$$J_p = \frac{1}{32}\pi d^4\left(1-\left(r/R\right)^4\right),$$

where d — an exterior string diameter.

The drag torque caused by forces of viscosity at rotary movement of a string in flush fluid, is calculated by the formula of Margules

$$\tau = \frac{4\pi\eta h\varepsilon^2}{(\varepsilon^2 - 1)r_T}\omega;\ \varepsilon = \frac{R_C}{r_T},$$

where η — a coefficient of viscosity for flush fluid; h — length of a lease of a string; R_C — a well radius; r_T — exterior radius of section of a drill-pipe.

Reproduction of interaction of a chisel with a stope is an important point in the course of constructing of computer model of a drill column. At cone rock chisel rotation there is a complex process of running of teeth of a frame set of a rolling cutter on a stope and their bulging-in in the formation, causing its breaking down. On fig. 3 a single act of running of a tooth of a conical cutter round instant centre of rotation B_1 (B_2) is schematically demonstrated. Thereof the cone axis migrates from dot O_1 in dot O_2, and the chisel body is moved in a vertical direction on magnitude

$$h = r_в\left(\cos\left(\frac{\pi}{z} - \omega_{rc}t\right) - \cos\frac{\pi}{z}\right)\sin\beta,$$

where ω_{rc} — angular speed of rolling cutter rotation; $r_в$ — radius of a peripheral frame set of a rolling cutter; z — number of teeth of a rolling cutter; β — an angle between a fulcrum of a chisel and cone axis.

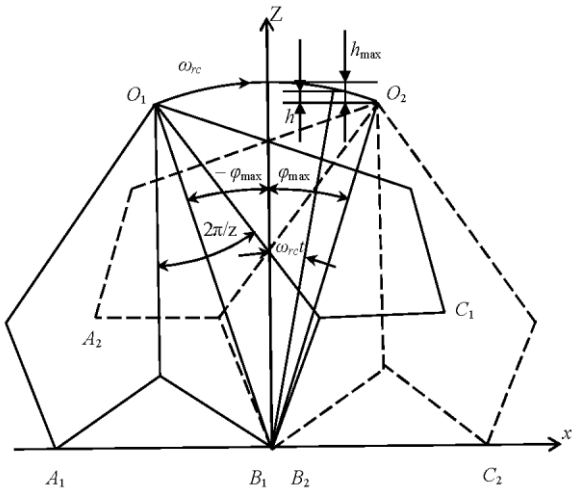


Fig. 3. Rotation of a peripheral frame set of a rolling cutter round instant centre of rotation

At rolling cutter rotation on an angle $\varphi = \omega_{rc}\ t$ there is a moment of force enclosed to cone axis

$$M_{rc} = r_в P\,\mathrm{tg}\left(\varphi_{rc} - \frac{\pi}{z}\right),$$

where P — the force operating on a chisel from a drill column. In position $\varphi = 0$ moment of force M_{rc} modifies a sign, i.e. on a gap $[-\varphi_{max}, 0]$ moment of force counteracts chisel rotary movement, and on a gap $[0, \varphi_{max}]$ speeds it up.

Feature of process of running of a rolling cutter is that time t and an angle $\varphi = \omega_{rc} t$ are in gaps accordingly $0 \le t \le T_1$ and $0 \le \omega_{rc} t \le \frac{2\pi}{z}$, where T_1 — time of running of a rolling cutter from tooth to tooth. Arguments h, T_1 are instituted proceeding from assumptions that all roller cones are synchronously rolled on a stope, i.e. simultaneously contact it by one or two teeth. At the moment of transfer from one bearing tooth of a rolling cutter velocity is almost instantaneously modified to another from the maximum negative value to the maximum positive value. During this moment axled speed-up theoretically aspires to infinity, and in practice, allowing for strains of members, has very big numerical value.

At study of interaction of a tooth of a rolling cutter with formation functional dependence of resistance force of formation P_z which is caused by tooth bulging-in on depth δ is used. On fig. 4 linearized diagramme of such dependence is shown, having four leases which characterize different stages of interaction: elastic deformation (piece OA); a plastic deformation (piece AB); breaking down (piece BC or BC_1); unloading (piece CD or C_1D). At achievement of a lease of a plastic deformation or breaking down the rolling cutter tooth leaves recess owing to what the string loses a part of the potential energy.

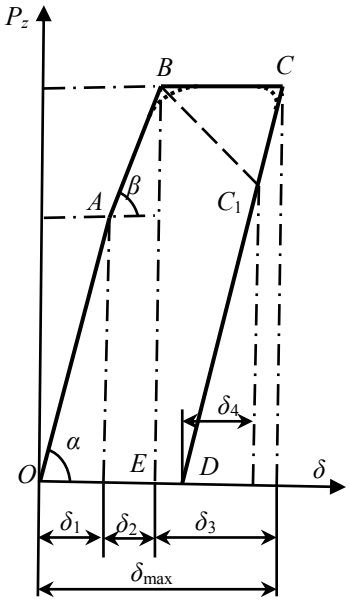


Fig. 4. A linearized diagram of elastoplastic strain of formation

For reproduction of dependencies which are linked to transmission of rotary movement of a chisel to rolling cutters, and also interactions of teeth with formation the subsystem represented on fig. 5 is used. It is necessary that at a bit complete revolution there is a uniform rock breaking without formation of a bottom-hole screw rack.

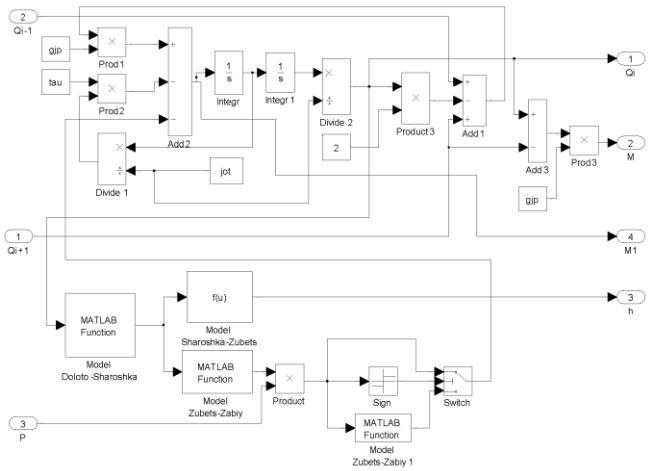


Fig. 5. A modulus of reproduction of bit interaction with formation

3. Results of Modelling

We use the obtained structural members for constructing of model of a drill column taking into account its longitudinal and rotary movement and interaction of a bit with a stope. As operating impact we will consider the force enclosed from a hoist through tackle system and moment of force from outside rotary of a table or a downhole motor. We will allow thus: heterogeneity of a drill column owing to use of polytypic drill-pipes; derrick strain at an offloading and its inertness; the resistance forces caused by interaction of flush fluid with walls of a string; floating force and flywheel action of a pole of flush fluid; resistance forces at interaction of a bit with formation. Source arguments of model are introduced in table 1 [2, 5].

Table 1. Key parameters of a mechanical part of the drilling unit

Argument	Value
Exterior diameter of a steel drill-pipes	127 mm
Inside diameter of a steel drill-pipes	113 mm
Length of a heavy weight drill-pipe	12.0 m
Exterior diameter of a heavy weight drill-pipes	178 mm
Inside diameter of a heavy weight drill-pipes	88 mm
Coefficient of resistance concerning longitudinal velocity of a drill-pipes	920 N·s/m
Gravity of matter of a drill-pipes	7990 kg/m^3
Elastic modulus for a drill-pipes material	$2.06 \cdot 10^{11}$ N/m^2
Gravity of flush fluid	1150 kg/m^3
Stiffness factor for a three spined up steel drill-pipes	$1.5141 \cdot 10^7$ N/m
Stiffness factor for a three spined up heavy weight drill-pipes	$3.2904 \cdot 10^8$ N/m
Mass of a tubular derrick	1.5 E$\cdot 10^4$ kg
Mass of flush fluid	6. $8271 \cdot 10^3$
Tubular derrick stiffness factor	1.02 E$\cdot 10^8$ N/m
Velocity of elastic deformation of a material of a drill-pipes	$5.1267 \cdot 10^3$ m/s

Argument	Value
Modulus of shearing	$8.0 \cdot 10^{10}$ N/m^2
Linear inertia moment of a steel drill-pipes	11.8628 kg·m
Polar inertia moment of a steel drill-pipes	$9.5325 \cdot 10^{-6}$ m^4
Linear inertia moment of a heavy weight drill-pipes	64.7377 kg·m
Polar inertia moment of a heavy weight drill-pipes	$9.2668 \cdot 10^{-5}$ m^4
Flush fluid coefficient of viscosity	0.014 Pa·s
Diameter of a chisel	190 mm
Diameter of rolling cutters	120 mm
Number of a teeth of a rolling cutter	24
Angle between a bit and rolling cutter fulcrum	52º

The drilling string is expedient for introducing in the form of multimass system with number of masses which corresponds to quantity of drilling candles. In process of string building-up the quantity of modules of model will be augmented. The structural simulink-model consists of same links, starting with the second one, which are devised in the form of a subsystem (fig. 1).

Links are also introduced into model for reproduction of dynamics of a derrick, a downhole motor, flywheel action of flush fluid in a hole, bit interaction with formation. Also in simulink-model there are modules for input impacts generation, and also registration and displaying of results.

By means of the obtained model the series of computing experiments has been conducted. On fig. 6 results of modelling of interaction of a cone rock bit with a hole bottom are displayed.

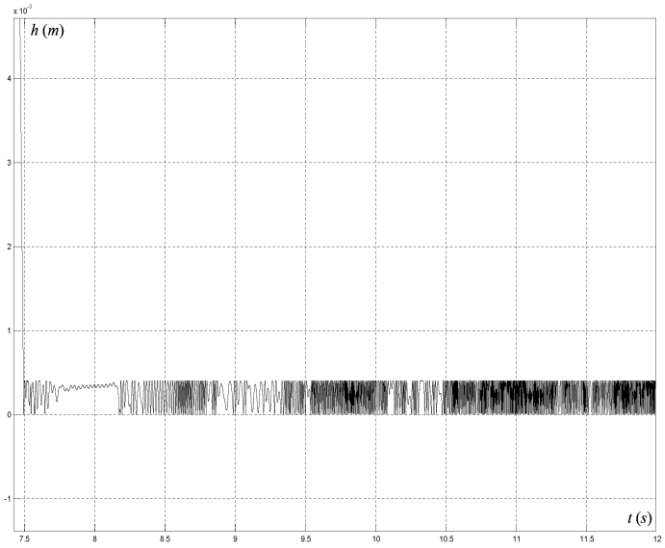


Fig. 6. A diagram of oscillatings of the body of bits

The string during the original moment of time was in a hover in flush fluid at altitude of 0,42 m from a hole bottom. For landing to a wall hook of tackle system the effort has been diminished by 720 kgf. At the moment of time 7,5 seconds there was a contact of roller cones to a hole bottom. On a drawing difficult nature of oscillatings of a bit is observed,

caused first of all by variations of dynamic effort from a string on a bit, and also the forces arising during flowing through of rolling cutters on a stope. Also the profile of variation of a bit weight (fig. 7) has been received at its contact to a hole bottom.

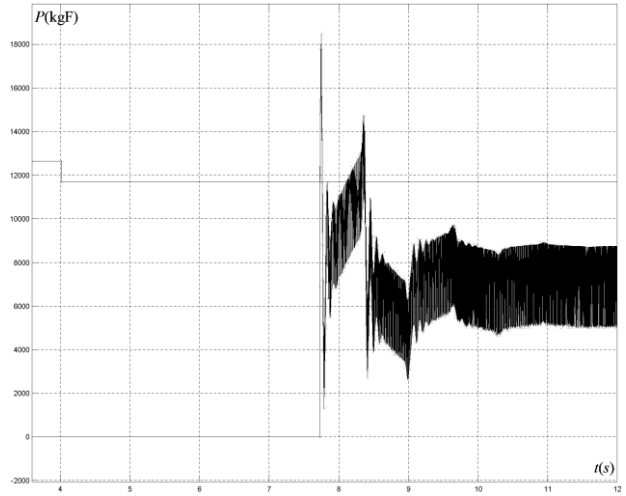


Fig. 7. A diagram of variation of a bit weight

4. Conclusion

Results of modelling testify that the devised structural model can be used with success both at a design stage of the equipment of the drilling unit, and at a drilling string assembly stage. Besides, the model can be used in a control system at a penetration of a hole for support of optimum drilling practicess, and also for optimization of trigger and elevating operations.

References

1. Balitsky P. V. Interplay of a drill column with a bottom hole. — Moscow: Mineral resources, 1975.
2. Abubakirov V. F. and Other. Drilling equipment: the Reference book. — T 1. — Moscow: Mineral resources, 2000.
3. Kiselyov N. V. and Other. Electric drives with the distributed arguments. — Leningrad: Shipbuilding, 1985.
4. Harchenko E. V. Dynamic processes of drilling units. — Lviv: Svit, 1991.
5. Abubakirov V. F. and Other. Drilling equipment: Reference book. — T 2. Ancillary. — Moscow: Mineral resources, 2003.
6. Navarro-López, E. Bit-sticking phenomena in a multi-degree-of-freedom controlled drillstring. Exploration and Production: Oil and Gas Review, 8(2), pp. 70-75, 2010.

METHODS OF CREATING MATHEMATICAL MODELS OF DYNAMIC LINKS WITH DISTRIBUTED PARAMETERS

M.M. Karimov, Sh. M. Gulyamov, M.M. Kadirov, S. M. Sagatova

Tashkent State Technology University, Uzbekistan,
informtgtu@mail.ru, mirhusan@mail.ru

Abstract. Methods for constructing mathematical models of dynamic objects and control. The features of the location of the transfer function of complex dynamic links with lumped and distributed parameters. Interpreted the line integral, the matrix and the method of functional transformations to find the transfer function of dynamic objects. Developed a method for modelling the structural and algorithmic processes in the electro - mechanical systems.

Keywords: dynamic links, the transfer functions of the system with distributed and lumped parameters, structural and algorithmic modelling electromechanical systems.

1. Introduction

Transfer functions are the traditional apparatus of representation models in problems of mechanics, electrodynamics, control systems and other systems approach and structural method to the study of technical systems is also often associated with the use of transfer functions. Modelling software contained in packages of applied programs focus on the operations of the transfer functions. To maintain the continuity of the decision of problems of modelling, synthesis, design, and also for use and development of possible programming environments expedient in cases when this is possible) use the traditional model of the transfer function.

2. Transfer functions of complex dynamic objects

Let's consider the particular location of the transfer function for dynamical systems with lumped and distributed parameters.

It is known that if the behaviour of some systems with lumped parameters, which has one entrance and exit, described by a linear differential equation

$$L_t y(t) = f(t), \tag{1}$$

where L_t - differential operator on variable t;

$$L_t = \sum_{k=0}^{n} a_k \frac{d^k}{dt^k} \ ; \tag{2}$$

y_t - reaction system on the input signal with zero initial conditions; $f(t)$ is the input signal, then the transfer function of the system has the form

$$W(p) = \frac{Y(p)}{F(p)}. \tag{3}$$

Here: $Y(p) = \tilde{L}\{y(t)\}$, $F(p) = \tilde{L}\{f(t)\}$, $\tilde{L}$ - operator of the Laplace transform.

If the system lumped has m input and n outputs, the transfer function of the object becomes the transfer matrix linking reaction of the object on each output with an effect on every login:

$$W(p) = \|W_{ik}\|,\ i = 1,2,\ldots,n;\ k = 1,2,\ldots,m\ . \tag{4}$$

If the system contains distributed parameters and is described by differential equations in partial derivatives, then for the full features of its transition properties, you already use a matrix with a finite number of rows and columns. Two-point function

$$R(x_1, x_2; p) = \frac{Y(x_2, p)}{F(x_1, p)},\ x_1 \in P,\ x_2 \in Q\ , \tag{5}$$

where P - is the set of points of input action; Q - the set of points at which the observed output signal $Y(x2, x1); F(x1, p)$ - is the Laplace transform of the time-from the output signal and the input impact [1].

3. Methods of finding the transfer functions

Depending on the type of structure for objects with distributed parameters, you can use the following methods of finding their transfer functions: direct, integral, method of functional transformation matrix.

The direct method is the application of a Laplace transformation to the partial differential equations that describes the object with distributed parameters [2]. The partial derivatives with respect to time t disappear in equation images are only partial derivatives with respect to the spatial variable x. This means that the equations in the space of states turns into an ordinary differential equation, the solution of which is simpler than the solution of the equation in partial derivatives. Boundary conditions for the original equation are moving in the initial conditions for the equations in the state space. When solving this equation with given boundary conditions are determined by the transfer function object with distributed parameters.

Direct method is the most convenient to obtain models of objects with distributed parameters, described by a wave equation, because during the transfer functions links prone to bending, it is necessary to solve a differential equation with respect to images of the fourth, and in some cases higher order.

Integral method is most useful when you need to find the transfer function of the object from the entrance to various points out e.g. the installation of devices for registration of parameters in different points of the investigated oscillating stand). For finding all of transfer functions enough to determine green's function for a given object and then, substituting the specific values of the coordinates of the relation ξ, to get the desired transfer function from the input points to the exit points [3].

The integrated method can be used for multidimensional objects - the only difference is that the corresponding differential operators on coordinates will be multidimensional. However, the task of constructing the green's function of such operators is significantly complicated, and therefore more effective methods are described below.

The mathematical apparatus, which has proved itself in the study of processes in systems with distributed parameters are the discrete functional transformations [3], which uses the discrete Fourier transform.

For determination of the transfer functions of complex tree-like systems containing objects with distributed parameters and elements with lumped parameters convenient to use the matrix solutions [4].

The main advantages of the matrix method lies in the possibility of obtaining the transfer functions of complex heterogeneous systems, containing links with the distributed and lumped parameters. This method is especially useful when considering the substantially heterogeneous objects with distributed parameters when they are describing, for example, torsion, longitudinal and lateral bending.

The transfer functions of the complex of core systems, it is convenient to use in applying the matrix continuation of the solution. For long многопролетных beams every flight can be considered as a homogeneous rod, which is described by a system of operator equations. In matrix notation this system gives the matrix of the span. Connection conditions spans allow to make a matrix transitions.

Complex three-dimensional spatial rod systems are reduced to the computing systems and next, calculated as a multi-dimensional beam or system of the beams. In this case, again get a matrix linking the two status points or planes of the system. Conditions for fastening the ends of the beams give additional equations and further definition of the transfer function of difficult rod systems is performed for both homogeneous rod.

4. Numerical simulation of dynamic objects with distributed parameters

The methods of forming of transfer functions allow to receive the equivalent model objects with distributed parameters in the form of transfer functions. Dynamic objects with distributed parameters describes the complex transfer functions, namely:

•fractional-rational transfer functions of high order,

$$W(p)=\frac{a_n p^n+\ldots+a_0}{b_{n+1}p^{n+1}+b_n p^n+\ldots+b_0},$$

where $n>20;$

•transcendental transfer functions (for example, $e^{-p},(ch(p))^{-1},(sh(p))^{-1},th(p),ctg(p))$;

•irrational transfer functions (in particular, $\frac{1}{\sqrt{p+1}},\frac{1}{\sqrt{p}},e^{-\sqrt{p}}$ etc).

From the analysis of the above mathematical models of electro-mechanical systems (EMS) with distributed parameters, it should be that the main feature of their transfer functions of Electromechanical systems is their transcendence. As a rule, the objects with distributed parameters describes the complex hyperbolic transfer functions type

$$W(p)=\frac{\sum_{i=0}^{n}A_i(p)\operatorname{ch}F(p)+\sum_{j=0}^{m}B_j(p)\operatorname{sh}F(p)}{\sum_{r=0}^{k}C_r(p)\operatorname{ch}F(p)+\sum_{q=0}^{\ell}D_q(p)\operatorname{sh}F(p)}, \quad (6)$$

where $A_i(p),A_j(p),C_r(p),D_q(p)$ - some rational, transcendental or irrational functions.

Numerical modeling of the rational way of presentation of models of objects with distributed parameters is with dependencies

$$\begin{aligned} \operatorname{ch} F(p) &= \frac{1}{2}\left(e^{F(p)}\left(1+e^{-2F(p)}\right)\right); \\ \operatorname{sh} F(p) &= \frac{1}{2}\left(e^{F(p)}\left(1-e^{-2F(p)}\right)\right), \end{aligned} \tag{7}$$

who is allowed to write (6) in the form:

$$W(p) = \frac{e^{-2F(p)}\left[\sum_{i=0}^{n} A_i(p) - \sum_{j=0}^{m} B_j(p)\right] + \sum_{i=0}^{n} A_i(p) + \sum_{j=0}^{m} B_j(p)}{e^{-2F(p)}\left[\sum_{r=0}^{k} C_r(p) - \sum_{q=0}^{\ell} D_q(p)\right] + \sum_{r=0}^{k} C_r(p) + \sum_{q=0}^{\ell} D_q(p)} =$$

$$= \frac{e^{-2F(p)}M(p) + N(p)}{e^{-2F(p)}P(p) + Q(p)}, \tag{8}$$

where

$$\left.\begin{aligned} M(p) &= \sum_{i=0}^{n} A_i(p) - \sum_{j=0}^{m} B_j(p); \\ N(p) &= \sum_{i=0}^{n} A_i(p) + \sum_{j=0}^{m} B_j(p); \\ P(p) &= \sum_{r=0}^{k} C_r(p) - \sum_{q=0}^{\ell} D_q(p); \\ Q(p) &= \sum_{r=0}^{k} C_r(p) + \sum_{q=0}^{\ell} D_q(p). \end{aligned}\right\}$$

Representation (8) lets you build your model, an advantage of which there are only two links with transfer functions .

Research in mathematical and computer modelling of complex heterogeneous systems, which include EMS, used in oil and gas, mining and transport industries, and construction of optimum control systems such objects over the last several decades are intensively conducted in practically all industrialized countries.

The same dynamic object analytically equivalent can be described by differential or integral equations, weight or transfer function, equation of state, etc. that leads to different structural-algorithmic variants of models. Under the structure is a set of blocks (action), distributed in space. Under algorithm is a sequence of actions over time. The most simple structure has the appearance of a serial connection of blocks. However, in the algorithms it is possible to perform some operations at the same time, which leads to the possibility of the organization of parallel branches, and in the structures of the possible transmission of the signals of some blocks sequentially, which gives the possibility of merging the blocks.

5. Conclusion

Thus, an additional a priori information can be taken into account with the help of the structural organization that improve the quality of the computing process and reduce the

solution time, improve accuracy or sustainability). It is obvious that the selection of parallel branches allows to reduce the solution time. Less obvious is the fact that the unification of units allows to increase the accuracy of the solution (for example, during the simulation of typical parts EMS [1]).

Method of structural design provides effective computer implementation of the model, taking into account the engineering requirements of the user requirements to the quality of the results, including consideration of any available additional a priori information about an object modelling.

A method is proposed structural-algorithmic modelling EMS, which represents an application of a set of models, implemented as a library of basic algorithms, obtained on the basis of a detailed mathematical description of the processes occurring in the sector of EMS. The main advantages of this method of modelling are that, first, it gives a graphic representation of arbitrarily complex system being modelled; second, it allows the same way to describe objects using models of any form, including impulse and transitional transfer functions, and the rules of connection of individual blocks; third, it provides the possibility to define the characteristics of the whole system and its individual parts, analyze and synthesize complex objects containing the links as with lumped and distributed parameters.

Reference:

1. Verlan A.F. Electronic modeling transfer functions / A.F. Verlan, V.F. Evdokimov. - K.: Tehnika, 1970. – 232p.
2. Doetsch G. Guide to the Practical Application of the Laplace transform and Z-transform / G. Doetsch. - Moscow: Nauka, 1971. - 288 p.
3. Rassudov L.N. Actuators distributed parameter mechanical elements / L.N. Rassudov, V.N. Myadzel. - Leningrad: Leningrad branch Energoatomizdat, 1987. – 143p.
4. Kiselev N.V. Actuators distributed parameter / N. Kiselev, V.N. Myadzel, L.N. Rassudov. - L.: Shipbuilding, 1985. – 220p.
5. Sadovoy A.V. Systems optimal control precision electric / Sadovoy A.V., B.V. Sukhinin, Y.V. Sohina - Kiev: ISIMO, 1996. - 298p.

Retrial Queueing System M|GI|1 Researching by Means of the Second-order Asymptotic Analysis Method under a Heavy Load Condition

Ekaterina Fedorova[1], Anatoly Nazarov[2]

[1,2] Tomsk State University, Russian Federation
[1] moiskate@mail.ru, [2] nazarov.tsu@gmail.com

Abstract. In the paper the retrial queuing system M|GI|1 is researched by means of the second-order asymptotic analysis method under heavy load condition. During the investigation, the direct system of Kolmogorov differential equations is made. In considering condition the asymptotic equations are derived and the asymptotic functions are given. Finally an asymptotic characteristic function of number of calls in the orbit is obtained.

Key words: retrial queueing systems, asymptotic analysis method, heavy load.

In queueing theory there are generally two classes of queueing systems: systems with queue and loss systems. In real systems there are occasions when queue cannot be explicitly identified, but also we cannot say that nonserviced call is lost if it comes when the service is impossible. Usually, the source of calls does not refuse to be serviced, and it performs repeated calls to get the desired service. An example of such systems is telecommunication systems.

In this regard, a new class of queuing systems has been provided – the systems with a source of repeated calls or retrial queueing systems [1, 2].

The majority of retrial queueing systems researching are carried out numerically or via imitate simulation. Analytical results are obtained only in cases simple input flow models and service discipline (eg. Poisson input process and the exponential distribution service law).

We propose a method of an asymptotic analysis under a heavy load condition for the single-line retrial queueing systems researching.

Mathematical model

Consider a single-line retrial queueing system, the input process is distributed by Poisson's law with parameter λ and the service time of each call has generally independent distribution function $B(x)$. If a call is received when a service device is free, it takes the device for maintenance. If the device is busy, the call goes to orbit, where it performs a random delay. A duration of that delay has an exponential distribution with parameter σ. From the orbit after a random delay, the call once again refers to the service unit to attempt to re-take it. If the device is free, the call is to serve by it, otherwise the application instantly returns to the source of repeated calls for the next delay implementation.

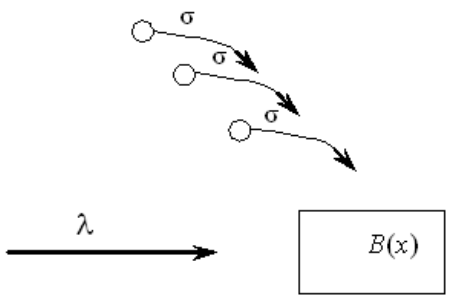


Fig. 1. Retrial queueing system M|GI|1

In a number of previous papers devoted to the study of various single-line retrial queueing system [3, 4], we demonstrated that this method has a fairly narrow range of applicability: for a load rate $\rho < 0.95$ Kolmogorov distance has values $\Delta > 0.05$. In this regard, we propose to increase the accuracy of the approximation by getting the 2nd order asymptotic formula.

Let $i(t)$ denotes calls number in orbit, $z(t)$ is the length of the interval from time t until the end of the call service and $k(t)$ determines the state of the service device as follows:

$$k(t)=\begin{cases}0, \text{ if device is free},\\ 1, \text{ if device is busy}.\end{cases}$$

We denote $P\{k(t)=0,\ i(t)=i\}=P(0,i,t)$ - the probability that the device is free at time t and there are i calls in the orbit; and $P\{k(t)=1,\ i(t)=i,\ z(t)<z\}=P(1,i,z,t)$ - the probability that the device is busy at time t, there are i calls in the orbit and the remaining service time is less than z.

Obviously the process $\{k(t),\ i(t),\ z(t)\}$ is Markovian with a variable number of components.

Then the problem is to define the probability distribution of the calls number in orbit in this system.

We make the direct system of Kolmogorov differential equations for the RQ-system states probability distribution $P(1,i,z,t)$, $P(0,i,t)$:

$$\begin{cases}\dfrac{\partial P(0,i,t)}{\partial t}=\dfrac{\partial P(1,i,0,t)}{\partial z}-(\lambda+i\sigma)P(0,i,t),\\ \dfrac{\partial P(1,i,z,t)}{\partial t}=\dfrac{\partial P(1,i,z,t)}{\partial z}-\dfrac{\partial P(1,i,0,t)}{\partial z}-\lambda P(1,i,z,t)+\\ +(i+1)\sigma\cdot P(0,i+1,t)B(z)+\lambda P(0,i,t)B(z)+\lambda P(1,i-1,z,t).\end{cases} \tag{1}$$

In the stationary form the system (1) becomes:

$$\begin{cases}\dfrac{\partial P(1,i,0)}{\partial z}-(\lambda+i\sigma)P(0,i)=0,\\ \dfrac{\partial P(1,i,z)}{\partial z}-\dfrac{\partial P(1,i,0)}{\partial z}-\lambda P(1,i,z)+(i+1)\sigma\cdot P(0,i+1)B(z)+\\ +\lambda P(0,i)B(z)+\lambda P(1,i-1,z)=0.\end{cases} \tag{2}$$

where $P(k,i,z,t)\equiv P(k,i,z)$.

In this system we turn to the characteristic functions: $H(0,u)=\sum_i e^{jui}P(0,i)$ and $H(1,u,z)=\sum_i e^{jui}P(1,i,z)$, where $j=\sqrt{-1}$ - the imaginary unit.

We introduce a parameter $\rho=\lambda b$ characterizing the system load, where b is a mean of the service time distribution. The stationary regime in such system exists for $\rho<1$.

Then the system (2) has the following form:

$$\begin{cases} b\dfrac{\partial H(1,u,0)}{\partial z}-\rho H(0,u)+j\sigma b\dfrac{\partial H(0,u)}{\partial u}=0, \\ b\dfrac{\partial H(1,u,z)}{\partial z}-b\dfrac{\partial H(1,u,0)}{\partial z}-e^{-ju}j\sigma b\dfrac{\partial H(0,u)}{\partial u}B(z)+\rho H(0,u)B(z)+ \\ +(e^{ju}-1)\rho H(1,u,z)=0. \end{cases} \quad (3)$$

Derivation of asymptotic equations

The system (3) will be solved by the method of asymptotic analysis under heavy load condition, that is $\lambda b=\rho\uparrow 1$, or when $\varepsilon\downarrow 0$, where $\varepsilon=1-\rho>0$ is an infinitesimal variable.

We introduce notations $u=\varepsilon w$, $H(0,u)=\varepsilon G(w,\varepsilon)$, $H(1,u,z)=F(w,z,\varepsilon)$. Then the system (3) can be rewritten as:

$$\begin{cases} b\dfrac{\partial F(w,0,\varepsilon)}{\partial z}-(1-\varepsilon)\varepsilon G(w,\varepsilon)+j\sigma b\dfrac{\partial G(w,\varepsilon)}{\partial w}=0, \\ b\dfrac{\partial F(w,z,\varepsilon)}{\partial z}-b\dfrac{\partial F(w,0,\varepsilon)}{\partial z}-e^{-jw\varepsilon}j\sigma b\dfrac{\partial G(w,\varepsilon)}{\partial w}B(z)+(1-\varepsilon)\varepsilon G(w,\varepsilon)B(z)+\cdot \\ +(e^{jw\varepsilon}-1)(1-\varepsilon)F(w,z,\varepsilon)=0. \end{cases} \quad (4)$$

From this system we derive asymptotic equations.

1. Denote $F(w,z)=\lim_{\varepsilon\to 0}F(w,z,\varepsilon)$ and $G(w)=\lim_{\varepsilon\to 0}G(w,\varepsilon)$. Then we have following system:

$$\begin{cases} \dfrac{\partial F(w,0)}{\partial z}=-j\sigma\dfrac{dG(w)}{dw}, \\ \dfrac{\partial F(w,z)}{\partial z}-\dfrac{\partial F(w,0)}{\partial z}(1-B(z))=0. \end{cases} \quad (5)$$

2. Let's write the expansions of functions in Taylor series:

$$G(w,\varepsilon)=G(w)+\varepsilon\cdot g(w)+\varepsilon^2\cdot\psi(w)+O(\varepsilon^3), \quad (6)$$

$$F(w,\varepsilon,z)=F(w,z)+\varepsilon\cdot f(w,z)+\varepsilon^2\cdot\varphi(w,z)+O(\varepsilon^3), \quad (7)$$

where $O(\varepsilon^3)$ - an infinitesimal of order ε^3.

Substituting formulas (6) and (7) in the system (4) and performing some actions on the equations, we obtain the following system:

$$\begin{cases} b\frac{\partial F(w,0)}{\partial z}+b\varepsilon\frac{\partial f(w,0)}{\partial z}+b\varepsilon^2\frac{\partial \varphi(w,0)}{\partial z}-(1-\varepsilon)\varepsilon G(w)-(1-\varepsilon)\varepsilon^2 g(w)- \\ -(1-\varepsilon)\varepsilon^3\psi(w)+j\sigma b\frac{dG(w)}{dw}+j\sigma b\varepsilon\frac{dg(w)}{dw}+j\sigma b\varepsilon^2\frac{d\psi(w)}{dw}=O(\varepsilon^3), \\ b\frac{\partial F(w,z)}{\partial z}+b\varepsilon\frac{\partial f(w,z)}{\partial z}+b\varepsilon^2\frac{\partial \varphi(w,z)}{\partial z}-b\frac{\partial F(w,0)}{\partial z}-b\varepsilon\frac{\partial f(w,0)}{\partial z}-b\varepsilon^2\frac{\partial \varphi(w,0)}{\partial z}- \\ -e^{-jw\varepsilon}j\sigma b\frac{dG(w)}{dw}B(z)-e^{-jw\varepsilon}j\sigma b\varepsilon\frac{\partial g(w)}{\partial w}B(z)-e^{-jw\varepsilon}j\sigma b\varepsilon^2\frac{\partial \psi(w)}{\partial w}B(z)+ \\ +(1-\varepsilon)\varepsilon G(w)B(z)+(1-\varepsilon)\varepsilon^2 g(w)B(z)+(1-\varepsilon)\varepsilon^3\psi(w)B(z)+ \\ +(e^{jw\varepsilon}-1)(1-\varepsilon)F(w,z)+(e^{jw\varepsilon}-1)(1-\varepsilon)\varepsilon f(w,z)+(e^{jw\varepsilon}-1)(1-\varepsilon)\varepsilon^2\varphi(w,z)=O(\varepsilon^3). \end{cases}$$

Rewrite the system for expressions with same powers of ε. We obtain 4 asymptotic equations more.

$$\begin{cases} b\frac{\partial f(w,0)}{\partial z}-G(w)+j\sigma b\frac{dg(w)}{dw}=0, \\ b\frac{\partial f(w,z)}{\partial z}-b\frac{\partial f(w,0)}{\partial z}+jwj\sigma b\frac{dG(w)}{dw}B(z)-j\sigma b\frac{dg(w)}{dw}B(z)+G(w)B(z)+jwF(w,z)=0, \\ b\frac{\partial \varphi(w,0)}{\partial z}+G(w)-g(w)+j\sigma b\frac{d\psi(w)}{dw}=0, \\ b\frac{\partial \varphi(w,z)}{\partial z}-b\frac{\partial \varphi(w,0)}{\partial z}-\frac{(jw)^2}{2}j\sigma b\frac{dG(w)}{dw}B(z)+jwj\sigma b\frac{dg(w)}{dw}B(z)-j\sigma b\frac{d\psi(w)}{dw}B(z)- \\ -G(w)B(z)+g(w)B(z)+(-jw+\frac{(jw)^2}{2})F(w,z)+jwf(w,z)=0. \end{cases} \quad (8)$$

3. Summing up the equations of the system (4), we have:

$$b\frac{\partial F(w,z,\varepsilon)}{\partial z}+(1-e^{-jw\varepsilon}B(z))j\sigma b\frac{\partial G(w,\varepsilon)}{\partial w}+(1-\varepsilon)\varepsilon G(w,\varepsilon)(B(z)-1)+$$
$$+(e^{jw\varepsilon}-1)(1-\varepsilon)F(w,z,\varepsilon)=0.$$

Rewrite this expression under condition that $z\to\infty$:

$$(1-e^{-jw\varepsilon})j\sigma b\frac{\partial G(w,\varepsilon)}{\partial w}+(e^{jw\varepsilon}-1)(1-\varepsilon)F(w,\varepsilon)=0.$$

Performing some actions and extracting the equations for expressions with same powers of ε, give following system:

$$\begin{cases} j\sigma b\frac{dG(w)}{dw}+F(w)=0, \\ j\sigma b\frac{dg(w)}{dw}+(jw-1)F(w)+f(w)=0, \\ j\sigma b\frac{d\psi(w)}{dw}+(-jw+\frac{(jw)^2}{2})F(w)+(-1+jw)\cdot f(w)+\cdot\varphi(w)=0. \end{cases} \quad (9)$$

When we join systems (5), (8) and (9), we obtain the system of seven asymptotic equations:

$$\begin{cases} b\dfrac{\partial F(w,0)}{\partial z}+j\sigma b\dfrac{dG(w)}{dw}=0, \\ \dfrac{\partial F(w,z)}{\partial z}-\dfrac{\partial F(w,0)}{\partial z}(1-B(z))=0, \\ b\dfrac{\partial f(w,0)}{\partial z}-G(w)+j\sigma b\dfrac{dg(w)}{dw}=0, \\ b\dfrac{\partial f(w,z)}{\partial z}-b\dfrac{\partial f(w,0)}{\partial z}+jwj\sigma b\dfrac{dG(w)}{dw}B(z)-j\sigma b\dfrac{dg(w)}{dw}B(z)+G(w)B(z)+jwF(w,z)=0, \\ b\dfrac{\partial \varphi(w,0)}{\partial z}+G(w)-g(w)+j\sigma b\dfrac{d\psi(w)}{dw}=0, \\ b\dfrac{\partial \varphi(w,z)}{\partial z}-b\dfrac{\partial \varphi(w,0)}{\partial z}-\dfrac{(jw)^2}{2}j\sigma b\dfrac{dG(w)}{dw}B(z)+jwj\sigma b\dfrac{dg(w)}{dw}B(z)-j\sigma b\dfrac{d\psi(w)}{dw}B(z)- \\ -G(w)B(z)+g(w)B(z)+(-jw+\dfrac{(jw)^2}{2})F(w,z)+jwf(w,z)=0. \\ j\sigma b\dfrac{dG(w)}{dw}+F(w)=0, \\ j\sigma b\dfrac{dg(w)}{dw}+(jw-1)F(w)+f(w)=0, \\ j\sigma b\dfrac{d\psi(w)}{dw}+(-jw+\dfrac{(jw)^2}{2})F(w)+(-1+jw)\cdot f(w)+\cdot\varphi(w)=0. \end{cases} \tag{10}$$

The beforelimited characteristic function $H(u)=H(1,u,\infty)+H(0,u)$ under heavy load conditions can be determined approximately by the equation:

$$H(u)\approx h(u)=F(w)+\varepsilon\{G(w)+f(w)\}.$$

Thus, to solve the problem it is necessary to solve the system (10) for unknowing functions $F(w)$, $G(w)$ and $f(w)$.

Investigation of the asymptotic equations system

1. Lets function $F(w,z)$ has multiplication form:

$$F(w,z)=R(z)\cdot\Phi(w). \tag{11}$$

It is obvious that $F(w)=\lim\limits_{z\to\infty}F(w,z)=\Phi(w)$.

By substituting formula (11) in the 2nd equation of the system (10) we obtain $R'(z)-R'(0)(1-B(z))=0.$ Solution of this equation is $R(z)=R'(0)\int\limits_0^z(1-B(x))dx.$ As $R(\infty)=1$, so $R'(0)\int\limits_0^\infty(1-B(x))dx=1.$ Thus we get $R'(0)=\dfrac{1}{b}$. In this way we have:

$$R(z)=\frac{1}{b}\int_0^z(1-B(x))dx. \tag{12}$$

The derivative of $G(w)$ can be expressed through function $\Phi(w)$ from the first equation of the system (10).

$$j\sigma\frac{dG(w)}{dw} = -\frac{1}{b}\Phi(w). \tag{13}$$

2. $\frac{dg(w)}{dw}$ is expressed from the 3rd equation of the system as follows.

$$j\sigma b\frac{dg(w)}{dw} = -b\frac{\partial f(w,0)}{\partial z} + G(w)\,. \tag{14}$$

Substitute the expression (14) in the 4th equation of system (10):

$$b\frac{\partial f(w,z)}{\partial z} - b\frac{\partial f(w,0)}{\partial z}(1-B(z)) + jwj\sigma b\frac{dG(w)}{dw}B(z) + jwF(w,z) = 0\,.$$

Then we insert here expressions (11) and (13) and integrate:

$$f(w,z) = \frac{\partial f(w,0)}{\partial z}\int_0^z (1-B(x))dx - \frac{jw}{b}\Phi(w)\int_0^z (R(x)-B(x))dx.$$

Considering formula (12) we can obtain

$$f(w,z) = \frac{\partial f(w,0)}{\partial z}bR(z) - \frac{jw}{b}\Phi(w)\int_0^z (R(x)-B(x))dx. \tag{15}$$

We rewrite this expression under limit $z \to \infty$:

$$f(w,\infty) = f(w) = \frac{\partial f(w,0)}{\partial z}b - \frac{jw}{b}\Phi(w)\int_0^\infty (R(x)-B(x))dx. \tag{16}$$

It is easy to show that $\int_0^\infty (R(x)-B(x))dx = b - \frac{1}{2b}b_2$, where b_2 is a second-order moment of the service time distribution law. Thus expression (16) can be rewritten as:

$$f(w) = \frac{\partial f(w,0)}{\partial z}b - jw\Phi(w)\left(1-\frac{b_2}{2b^2}\right). \tag{17}$$

Also following formula can be given:

$$b\frac{\partial f(w,0)}{\partial z} = f(w) + jw\Phi(w)\left(1-\frac{b_2}{2b^2}\right). \tag{18}$$

From equation (14) we have

$$j\sigma b\frac{dg(w)}{dw} = G(w) - f(w) - jw\Phi(w)\left(1-\frac{b_2}{2b^2}\right). \tag{19}$$

By substituting of formula (19) in the 8th equation of system (10) we get:

$$G(w) - jw\Phi(w)\left(1-\frac{b_2}{2b^2}\right) + (jw-1)\Phi(w) = 0\,.$$

We differentiate the equation:

$$G'(w) - jw\Phi'(w)\left(1-\frac{b_2}{2b^2}\right) - j\Phi(w)\left(1-\frac{b_2}{2b^2}\right) + (jw-1)\Phi'(w) + j\Phi(w) = 0\,.$$

Considering knowing functions we obtain:

$$\Phi'(w)\left[jw - \frac{2b^2}{b_2}\right] + j\Phi(w)\left[1 + \frac{2b}{\sigma b_2}\right] = 0\,.$$

Denote $\alpha = 1 + \frac{2b}{\sigma b_2}$, $\beta = \frac{2b^2}{b_2}$, then the solution is $\Phi(w) = \left(1 - \frac{jw}{\beta}\right)^{-\alpha}$.

3. The following expression can be obtained from the 5th equation of system (10):

$$j\sigma b \frac{d\psi(w)}{dw} = g(w) - G(w) - b\frac{\partial\varphi(w,0)}{\partial z}. \tag{20}$$

Substituting formula (20) in the 6th equation of system (10) gives:

$$b\frac{\partial\varphi(w,z)}{\partial z} - b\frac{\partial\varphi(w,0)}{\partial z}(1 - B(z)) - \frac{(jw)^2}{2} j\sigma b \frac{dG(w)}{dw} B(z) + jwj\sigma b \frac{dg(w)}{dw} B(z) -$$
$$+ (-jw + \frac{(jw)^2}{2}) F(w,z) + jwf(w,z) = 0.$$

Using the expressions (13), (16), (18) and (19) we can simply obtain the following formula:

$$\varphi(w,z) = \frac{\partial\varphi(w,0)}{\partial z}\int_0^z (1 - B(x))dx - \frac{jw}{b} G(w)\int_0^z B(x)dx + \frac{jw}{b} f(w)\int_0^z [B(y) - R(y)]dy +$$
$$+ \frac{(jw)^2}{b}\Phi(w)\left(-\frac{1}{2} + 1 - \frac{b_2}{2b^2}\right)\int_0^z B(x)dx + \frac{1}{b}(jw - \frac{(jw)^2}{2})\Phi(w)\int_0^z R(x)dx - \tag{21}$$
$$+ \frac{(jw)^2}{b^2}\Phi(w)\int_o^z \left\{-\left(1 - \frac{b_2}{2b^2}\right)R(y)b + \int_0^y (R(x) - B(x))dx\right\}dy.$$

From equation (13) it is known that

$$G(w) = \left(1 - \frac{jw}{\beta}\right)\Phi(w). \tag{22}$$

In this way, the formula (21) can be simplified. In case $z \to \infty$ we have

$$\varphi(w,\infty) = \varphi(w) = \frac{\partial\varphi(w,0)}{\partial z} b + jw\Phi(w)\left(1 - \frac{b_2}{2b^2}\right) - jwf(w)\left(1 - \frac{b_2}{2b^2}\right) +$$
$$+ (jw)^2\Phi(w)\left[\frac{1}{2}\left(1 - \frac{b_2}{2b^2}\right) + \frac{1}{b^2}\int_o^\infty \left\{\int_0^y (R(x) - B(x))dx - \left(b - \frac{b_2}{2b}\right)R(y)\right\}dy\right].$$

It is easy to show that the integral is equal:

$$\int_o^\infty \left\{\int_0^y (R(x) - B(x))dx - \left(b - \frac{b_2}{2b}\right)R(y)\right\}dy = \frac{1}{2}\left[\int_o^\infty y^2 dR(y) - b_2 - \left(b - \frac{b_2}{2b}\right)\int_o^\infty y^2 dR'(y)\right] = \frac{b_3}{6b} - \frac{b_2^2}{4b^2}.$$

So we obtain final formula for function $\varphi(w)$:

$$\varphi(w) = \frac{\partial\varphi(w,0)}{\partial z} b + jw\Phi(w)\left(1 - \frac{b_2}{2b^2}\right) - jwf(w)\left(1 - \frac{b_2}{2b^2}\right) + \tag{23}$$
$$+ (jw)^2\Phi(w)\left[\frac{1}{2} - \frac{b_2}{4b^2} + \frac{b_3}{6b^3} - \frac{b_2^2}{4b^4}\right].$$

4. Substituting formulas (20) and (23) in the last equation of system (10) gives:

$$g(w) = G(w) + \Phi(w)\cdot\left(jw\frac{b_2}{2b^2} - (jw)^2\left[1 - \frac{b_2}{4b^2} + \frac{b_3}{6b^3} - \frac{b_2^2}{4b^4}\right]\right) + f(w)\left(1 - jw\frac{b_2}{2b^2}\right).$$

This equation is differentiated and equated with expression (19):

$$j\sigma b f'(w)\left(1-jw\frac{b_2}{2b^2}\right)+f(w)\frac{\sigma b_2}{2b}+f(w)=G(w)-jw\Phi(w)\left(1-\frac{b_2}{2b^2}\right)-j\sigma bG'(w)-$$

$$-j\sigma b\Phi'(w)\cdot\left(jw\frac{b_2}{2b^2}-(jw)^2\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right)-j\sigma b\Phi(w)\cdot\left(j\frac{b_2}{2b^2}+2w\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right).$$

The equation can be simplified by considering formula (22) and $\Phi'(w)=\frac{j\alpha}{\beta}\left(1-\frac{jw}{\beta}\right)^{-1}\Phi(w)$. Thus we have

$$f'(w)\cdot j\sigma b\left(1-jw\frac{b_2}{2b^2}\right)+f(w)\left(1+\frac{\sigma b_2}{2b}\right)=\Phi(w)\left[2+\sigma b\frac{b_2}{2b^2}-\right.$$
$$\left.-jw\left(1+2\sigma b\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right)+jw\frac{\alpha\sigma b}{(\beta-jw)}\cdot\left(\frac{b_2}{2b^2}-jw\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right)\right] \quad (24)$$

We solve the differential equation (24) by the method of variation of parameters. So the solution has the following form :

$$f(w)=C(w)\cdot\left(1-jw\frac{b_2}{2b^2}\right)^{-\left(1+\frac{2b}{\sigma b_2}\right)}=C(w)\cdot\Phi(w)$$

From the equation (24) the expression for $C(w)$ can be obtained:

$$C(w)=\int_0^w\left(1-\frac{jy}{\beta}\right)^{-1}\times\left[\frac{2}{j\sigma b}+\frac{b_2}{2jb^2}-\frac{y}{\sigma b}\left(1+2\sigma b\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right)+\right.$$
$$\left.+y\frac{\alpha}{\beta}\left(1-\frac{jy}{\beta}\right)^{-1}\cdot\left(\frac{b_2}{2b^2}-jy\left[1-\frac{b_2}{4b^2}+\frac{b_3}{6b^3}-\frac{b_2^{\,2}}{4b^4}\right]\right)\right]dy.$$

Returning to the variable $u=\varepsilon w$ and parameter ρ. The characteristic function is following:

$$h(u)=F\left(\frac{u}{1-\rho}\right)+(1-\rho)\{G\left(\frac{u}{1-\rho}\right)+f\left(\frac{u}{1-\rho}\right)\}=$$
$$=\left(1-\frac{ju}{(1-\rho)\beta}\right)^{-\alpha}\left\{1+(1-\rho)\left(\left(1-\frac{ju}{(1-\rho)\beta}\right)^{-1}+C\left(\frac{u}{1-\rho}\right)\right)\right\}$$

Conclusion

During the researching, an asymptotic characteristic function of number of calls in the orbit is obtained. A numerically comparison of asymptotic and pre-limit characteristic function shows that the range of the method applicability is increased by 4 times: for load rate $\rho < 0.8$ Kolmogorov distance has values $\Delta \le 0.05$. In this regard the second-order asymptotic analysis method under a heavy load condition can be applied also for retrial queueing systems with no Poisson input process (eg. MMPP-process, SM-process), where pre-limit characteristic function is not obtained.

References

1. Falin G.I, Templeton J.G.C. Retrial queues. London: Chapman & Hall, 1997.
2. Artolejo J.R., Gomez-Corral A. Retrial Queueing Systems: A Computational Approach. – Springer, 2008.
3. Nazarov A.A., Moiseeva E.A. Issledovanie RQ-systemy M|GI|1 metodom asymptoticheskogo analiza v uslovii bolshoy zagruzki // Materialy mezhdunarodnoy nauchnoy konferencii «Sovremennye veroztnostnye metody analiza, proektirovania I optimizacii informacionno-telekommunikationnyh setey », 22. – Minsk: Izdatelskiy centr BGU, 2013. – P. 106-113 (in Russian).
4. Nazarov A.A., Moiseeva E.A. Issledovanie RQ-systemy MMPP|M|1 methodom asymptoticheskogo analiza v uslovii bolshoy zagruzki // Izvestia TPU. Matematica I Mekhanika. Fizika., 322, 2. - Tomsk: Izdatelstvo TPU, 2013. - P. 19-23 (in Russian).

Calculation of the Probability that a Gaussian Vector Falls in the Hyperellipsoid with the Uniform Density

Anatoly Nazarov[1], Alexander Moiseev[2]

[1,2] Tomsk State University, Russian Federation
[1] nazarov.tsu@gmail.com, [2] alexander-moiseev@mail.ru

Abstract. An investigation of falls of the Gaussian vector with correlated entries into hyperellipsoid with the uniform density is considered in the paper. A problem of calculation of the probability of the falls is solved, and some related problems solutions are presented in the paper also.

Key words: probability theory, Gaussian vector, uniform density hyperellipsoid.

Introduction

The probability that a Gaussian vector with independent entries falls in the hyperellipsoid with uniform density was found in the paper [1]. In this paper we solve this problem for a Gaussian vector with correlated entries. In addition to this, we solve an inverse problem – calculation of the hyperellipsoid size with given fall probability. A problem of calculation of maximum values for each coordinate variable of such hyperellipsoid is solved here too.

Presented research results are important for solving of some technical problems. As an example it can be the following problem. Consider a system of distributed data processing [2]. It is known [3] that in such system under a condition of high-rate message input [4], the distribution of processing messages in each system's node can be approximated by multi-dimensional Gaussian distribution. It is important to know a minimal number of processing devices in each node which make sure that the messages lost rate will be bounded by given small value. So, we can use results of the paper to calculate those numbers as maximum values for each coordinate variable of hyperellipsoid with given uniform density for Gaussian vector which parameters calculated by formulas from paper [3].

Let's $\boldsymbol{\xi}^{\mathrm{T}} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ is a Gaussian vector which has correlated entries, means equal to zeros and a covariance matrix $\mathbf{K} = \mathrm{M}\{\boldsymbol{\xi} \cdot \boldsymbol{\xi}^{\mathrm{T}}\}$. Denote by $\mathbf{x}^{\mathrm{T}} = \{x_1, x_2, \ldots, x_n\}$ vector of variables x_k for $k = \overline{1,n}$, and by $F_n(r)$ a probability that random vector ξ falls into n-dimensional hyperellipsoid $\mathrm{B}_n(r)$ with uniform density which is defined by the equation

$$\mathbf{x}^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{x} = r^2, \tag{1}$$

Here, r is some parameter which can be a characteristic of the hyperellipsoid size. So, we can write the following expression

$$F_n(r) = \mathrm{P}\{\boldsymbol{\xi} \in \mathrm{B}_n(r)\}. \tag{2}$$

The goal of the research is determination of the probability function $F_n(r)$ for given covariance matrix **K**.

Finding of the function $F_n(r)$

Denote by **A** a matrix which satisfies a condition that entries of the Gaussian vector

$$\boldsymbol{\eta} = \mathbf{A} \cdot \boldsymbol{\xi} \tag{3}$$

are independent standard normal random variables (they all has zero means and unit variances). Covariance matrix of the random vector $\boldsymbol{\eta}$ is diagonal identity. Denote it by **I**. So, the following statement is fulfilled

$$\mathbf{I} = \mathrm{M}\{\boldsymbol{\eta} \cdot \boldsymbol{\eta}^{\mathrm{T}}\} = \mathrm{M}\{\mathbf{A}\boldsymbol{\xi}\, \boldsymbol{\xi}^{\mathrm{T}}\mathbf{A}\} = \mathbf{A}\mathbf{K}\mathbf{A}^{\mathrm{T}},$$

and we rewrite it as follows:

$$\mathbf{A}\mathbf{K}\mathbf{A}^{\mathrm{T}} = \mathbf{I}. \tag{4}$$

Let's vector **y** is defined by statement

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{5}$$

and so,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}. \tag{6}$$

We find an equation for surface $\Lambda_n(r)$ which is a reflection (5) of the hyperellipsoid $\mathrm{B}_n(r)$.

Using (1), we can write

$$\begin{aligned} \{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x},\, \mathbf{x}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{x} = r^2\} &= \left\{\mathbf{y} : \left(\mathbf{A}^{-1}\mathbf{y}\right)^{\mathrm{T}}\mathbf{K}^{-1}\left(\mathbf{A}^{-1}\mathbf{y}\right) = r^2\right\} = \\ &= \left\{\mathbf{y} : \mathbf{y}^{\mathrm{T}}\left(\mathbf{A}^{-1}\right)^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{A}^{-1}\mathbf{y} = r^2\right\} = \{\mathbf{y} : \mathbf{y}^{\mathrm{T}}\mathbf{H}\mathbf{y} = r^2\}, \end{aligned} \tag{7}$$

where $\mathbf{H} = \left(\mathbf{A}^{-1}\right)^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{A}^{-1}$. Let's find an expression for the matrix **H**. To do that, consider its inversion:

$$\mathbf{H}^{-1} = \left[\left(\mathbf{A}^{-1}\right)^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{A}^{-1}\right]^{-1} = \mathbf{A}\mathbf{K}\mathbf{A}^{\mathrm{T}} = \mathbf{I}.$$

And so, matrix H is identity matrix. Now we can rewrite expression (7) as follows:

$$\{\mathbf{y} : \mathbf{y} = \mathbf{A}\mathbf{x},\, \mathbf{x}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{x} = r^2\} = \{\mathbf{y} : \mathbf{y}^{\mathrm{T}}\mathbf{y} = r^2\}.$$

In this manner, the expression (5) transforms hyperellipsoid $\mathrm{B}_n(r)$ into hyperball $\Lambda_n(r)$ which has a radius equal to *r* and the following surface equation

$$\mathbf{y}^{\mathrm{T}}\mathbf{y} = r^2 .$$

By reason of the expressions (2) and (3) we can write

$$F_n(r) = \mathrm{P}\{\boldsymbol{\xi} \in \mathrm{B}_n(r)\} = \mathrm{P}\{\boldsymbol{\eta} \in \Lambda_n(r)\}. \tag{8}$$

So, $F_n(r)$ is a distribution function of the argument r for the length $\zeta = \sqrt{\sum_{k=1}^{n} \eta_k^2}$ of the Gaussian vector $\boldsymbol{\eta}$ with independent standard entries.

It's known that for $n = 2$ the distribution (8) is a Rayleigh distribution

$$F_2(r) = 1 - \exp\left\{-\frac{r^2}{2}\right\},\ r \geq 0$$

and for $n = 3$ it is a Maxwell distribution

$$F_3(r) = 2\Phi(r) - \sqrt{\frac{2}{\pi}} \exp\left\{-\frac{r^2}{2}\right\},\ r \geq 0 .$$

Let's find a form of the function $F_n(r)$ for $n \geq 4$.

Probability density function $g(y_1, y_2, \ldots, y_n)$ of the vector $\boldsymbol{\eta}$ has a form

$$g(y_1, y_2, \ldots, y_n) = \prod_{k=1}^{n} \left(\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y_k^2}{2}\right\} \right),$$

and so, from expression (8), it follows that

$$F_n(r) = \mathrm{P}\{\boldsymbol{\eta} \in \Lambda_n(r)\} = \iint\limits_{\Lambda_n(r)} \ldots \int g(y_1, y_2, \ldots, y_n)\, dy_1\, dy_2 \ldots dy_n =$$

$$= (2\pi)^{-\frac{n}{2}} \iint\limits_{\sum_{k=1}^{n} y_k^2 \leq r^2} \ldots \int \exp\left\{-\frac{1}{2}\sum_{k=1}^{n} y_k^2\right\} dy_1\, dy_2 \ldots dy_n . \tag{9}$$

Let's transform a subintegral expression (9) into polar coordinates by using the substitutions

$$y_k = z v_k(\varphi_1, \varphi_2, \ldots, \varphi_{n-1}),\ k = \overline{1, n}, \tag{10}$$

where $z = \sqrt{\sum_{k=1}^{n} y_k^2}$ and $v_k(\varphi_1, \varphi_2, \ldots, \varphi_{n-1})$ are some functions of the coordinate axes rotation angles $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$.

Let's write a Jacobian determinant of the transform (10) in the following way

$$J=\begin{vmatrix} \frac{\partial y_1}{\partial z} & \frac{\partial y_1}{\partial \varphi_1} & \dots & \frac{\partial y_1}{\partial \varphi_{n-1}} \\ \frac{\partial y_2}{\partial z} & \frac{\partial y_2}{\partial \varphi_1} & \dots & \frac{\partial y_2}{\partial \varphi_{n-1}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial y_n}{\partial z} & \frac{\partial y_n}{\partial \varphi_1} & \dots & \frac{\partial y_n}{\partial \varphi_{n-1}} \end{vmatrix} = \begin{vmatrix} v_1 & z\frac{\partial v_1}{\partial \varphi_1} & \dots & z\frac{\partial v_1}{\partial \varphi_{n-1}} \\ v_2 & z\frac{\partial v_2}{\partial \varphi_1} & \dots & z\frac{\partial v_2}{\partial \varphi_{n-1}} \\ \dots & \dots & \dots & \dots \\ v_n & z\frac{\partial v_n}{\partial \varphi_1} & \dots & z\frac{\partial v_n}{\partial \varphi_{n-1}} \end{vmatrix} = z^{n-1} w(\varphi_1, \varphi_2, \dots, \varphi_{n-1}),$$

where $w(\varphi_1, \varphi_2, \dots, \varphi_{n-1})$ is some function of the angles φ_k for $k = \overline{1, n-1}$. Using substitution (10), the integral (9) can be rewrite as the follows

$$F_n(r) = \int_0^r z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz \; (2\pi)^{-\frac{n}{2}} \iint_D \dots \int w(\varphi_1, \varphi_2, \dots, \varphi_{n-1})\, d\varphi_1\, d\varphi_2 \dots d\varphi_{n-1} = \\ = b_n \int_0^r z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz, \tag{11}$$

where $b_n = (2\pi)^{-\frac{n}{2}} \iint_D \dots \int w(\varphi_1, \varphi_2, \dots, \varphi_{n-1})\, d\varphi_1\, d\varphi_2 \dots d\varphi_{n-1}$ and D is the $(n-1)$-dimensional region of the rotation angles φ_k values range.

A value of the b_n is not explicit calculated, but we will find its value from expression (11) by using a normalization requirement of the function $F_n(r)$:

$$1 = F_n(\infty) = b_n \int_0^\infty z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz.$$

So,

$$b_n = \left(\int_0^\infty z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz \right)^{-1},$$

and the function $F_n(r)$ can be written as follows

$$F_n(r) = \left(\int_0^\infty z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz \right)^{-1} \cdot \int_0^r z^{n-1} \exp\left\{-\frac{z^2}{2}\right\} dz.$$

Using here a substitution $z = \sqrt{2t}$ we obtain the following expression

$$F_n(r) = \left(\int_0^\infty (2t)^{\frac{n-1}{2}} e^{-t} \sqrt{2} \frac{dt}{2\sqrt{t}} \right)^{-1} \cdot \int_0^{\frac{r^2}{2}} (2t)^{\frac{n-1}{2}} e^{-t} \sqrt{2} \frac{dt}{2\sqrt{t}} = \tag{12}$$

$$= \left(\int_0^{\infty} t^{\frac{n}{2}-1} e^{-t}\, dt \right)^{-1} \cdot \int_0^{\frac{r^2}{2}} t^{\frac{n}{2}-1} e^{-t}\, dt = \left[\Gamma\left(\frac{n}{2}\right) \right]^{-1} \Gamma_{r^2/2}\left(\frac{n}{2}\right),$$

where $\Gamma_{r^2/2}\left(\frac{n}{2}\right)$ is incomplete gamma function [5].

Expression (12) allows calculating a probability that the Gaussian vector ξ falls in the hyperellipsoid $B_n(r)$ for given n and r. But the inverse problem, calculating of value r for given probability, is interesting too. This problem can be solved by finding a solution of the equation (12) in suppose that r is unknown but the fall probability $F_n(r)$ is given. Denote by Π a probability that the vector ξ misses the hyperellipsoid $B_n(r)$. Values of r for different values of Π and n is presented in the table 1 and figure 1.

Table 1. Values of r for the hyperellipsoid dimensions n = 5, 10, 50, 100 and different values of the fall probability (1 – Π).

Hyperellipsoid dimension n	Fall probability (1 – Π)			
	$1 - 10^{-3}$	$1 - 10^{-6}$	$1 - 10^{-9}$	$1 - 10^{-12}$
5	4,529	5,991	7,120	8,077
10	5,440	6,846	7,934	8,858
50	9,309	10,612	11,613	12,463
100	12,225	13,495	14,468	15,291

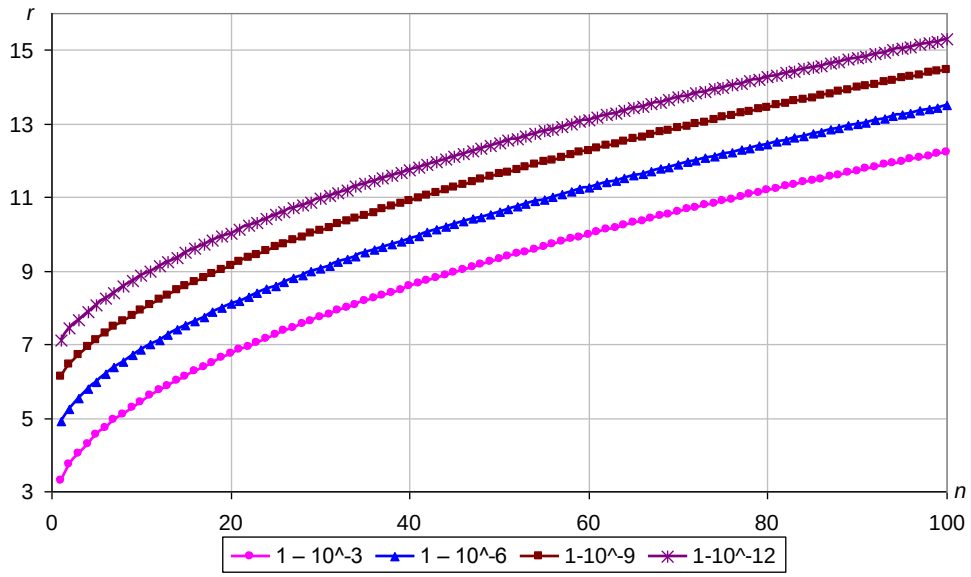


Figure 1. Value r dependence of the dimension n for different values of the fall probability (1 – П).

A problem of calculating of the boundary values of coordinate variables

Under a condition when a value of r and a correlation matrix are given, it is interesting to calculate boundary (maximum) values X_k of each coordinate variable x_k which satisfies the condition (2). This problem can be solved as the following constraint extremum finding problem

$$X_k = \max x_k ,$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{x} = r^2 .$$

Let's construct the following Lagrange function

$$L(x_1, x_2, \ldots, x_n, \lambda) = x_k + \lambda\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\left(K^{-1}\right)_{ij} x_j - r^2\right\} .$$

Necessary conditions of the constraint extremum are defined by the following system

$$\frac{\partial L(x_1, x_2, \ldots, x_n, \lambda)}{\partial x_i} = 2\lambda\sum_{j=1}^{n}\left(K^{-1}\right)_{ij} x_j = 0 \text{ when } i \neq k , \tag{13}$$

$$\frac{\partial L(x_1, x_2, \ldots, x_n, \lambda)}{\partial x_k} = 1 + 2\lambda\sum_{j=1}^{n}\left(K^{-1}\right)_{kj} x_j = 0 \text{ when } i = k , \tag{14}$$

$$\frac{\partial L(x_1, x_2, \ldots, x_n, \lambda)}{\partial \lambda} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \left(K^{-1}\right)_{ij} x_j - r^2 = 0 . \tag{15}$$

Denote an *i*-th row of matrix $\mathbf{K}^{-1}$ by $\left(\mathbf{K}^{-1}\right)^{<i>}$. The system (13)–(15) will be rewritten in the following matrix form

$$\left(\mathbf{K}^{-1}\right)^{<i>} \mathbf{x} = 0 \text{ when } i \neq k , \tag{16}$$

$$\left(\mathbf{K}^{-1}\right)^{<k>} \mathbf{x} = -\frac{1}{2\lambda} \text{ when } i = k , \tag{17}$$

$$\mathbf{x}^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{x} = r^2 . \tag{18}$$

Let's denote by $\mathbf{V}$ a vector with zero entries except an entry number *k* which is equal to $V_k = -1/2\lambda$. So, the system (16)–(17) will be rewritten as follows

$$\mathbf{K}^{-1} \mathbf{x} = \mathbf{V} .$$

And finally we obtain an expression

$$\mathbf{x} = \mathbf{K}\mathbf{V} . \tag{19}$$

Substituting this into (18), we obtain

$$\mathbf{V}^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{K} \mathbf{V} = r^2 ,$$

$$r^2 = \mathbf{V}^{\mathrm{T}} \mathbf{K}^{\mathrm{T}} \mathbf{V} = \frac{1}{(2\lambda)^2} K_{kk} ,$$

and so,

$$\frac{1}{2\lambda} = \pm \sqrt{\frac{r^2}{K_{kk}}} .$$

Using this fact, from the equation (19) we can obtain the following expression for the *k*-th entry of the vector $\mathbf{x}$:

$$x_k = -\frac{1}{2\lambda} K_{kk} = r\sqrt{K_{kk}} .$$

So, it is shown that the maximum value X_k of each coordinate variable x_k which satisfies a condition (2) can be calculated by the formula

$$X_k = r\sqrt{K_{kk}} . \tag{20}$$

Conclusion

A solution of the problem of calculating of the probability that Gaussian vector falls in the hyperellipsoid with uniform density is presented in the paper. The expression (12) for calculating this probability is obtained. Additionally, expression (12) can be used as an

equation for solving the inverse problem – finding value of hyperellipsoid size r when a fall probability is given.

Expression (20) can be used for calculating maximum values of coordinate variables for the uniform density hyperellipsoid when a fall probability is given. Such problem arrives in many tasks of technical systems design, for example during solving the problem of calculating of minimum number of processing devices that guarantees fixed level of the data lost probability for the system described in the paper [2].

References

1. Wentzel E.S., Ovcharov L.A. Probability theory and its engineering applications. Moscow, Nauka. 1988 (in Russian).
2. Grachev V.V., Moiseev A.N., Nazarov A.A., Yampolsky V.Z. Multi-stage queueing model of the distributed data processing system // Proceedings of TUSUR, 2012, No. 2, part 2, P. 248–251 (in Russian).
3. Nazarov A.A., Moiseev A.N. Analysis of an open non-Markovian $GI-(GI|\infty)^K$ queueing network with high-rate renewal arrival process // Problems of Information Transmission, 2013, Vol. 49, No. 2, P. 167–178.
4. Moiseev A., Nazarov A. Investigation of high intensive general flow // Proc. of the IV International Conference “Problems of Cybernetics and Informatics” (PCI’2012), Sept. 12-14, 2012 Baku, Azerbaijan. Baku: IEEE, P. 161–163.
5. Abramowitz M., Stegun I. Handbook of mathematical functions. Dover Publications, 1964.

Agent-Oriented View to Constriction of Innovation Environment: Technology Transfer Office

Rumen Andreev[1], Tsvetelina Yorgova[2]

[1] BAS, 1113 Sofia, Bulgaria
rumen@isdir.bas.bg

[2] ISSK- BAS; 1113 Sofia, Bulgaria
kostadinov@imbm.bas.bg

Absract. The paper reviews consists an innovation system of two layers – process-oriented and service-oriented. The process - oriented layer is responsible for carrying out of an innovation process. Its development reflects the way of performance of an innovation. If it ensures the realization of specific innovation style, it is regarding as *innovation platform* that is the foreground of the innovation system. The service-oriented layer supports the occurrence of an innovation. It is considered as system's background, i.e. *innovation environment*.

Keywords. Innovation process, innovation platform, innovation environment

1. Introduction

To cope with the development of an innovation environment, we investigate the applicability of multi-agent technology in its design and implementation. Hence, it is necessary to analyse this environment from the viewpoint of the multi-agent paradigm and its fundamental properties. The latter are predetermined by the nature of the entities used in an agent technology. The main characteristics of an agent, which development is at different stages in different agent systems are the followings [1, 2]:

- It is an entity with well-defined boundaries and interfaces, i.e. clearly identifiable;
- It solves problems and could be considered as an expert (problem solver);
- It is autonomous, i.e. self-controlled, self-organized;
- It performs a specific role (specific tasks);
- It exists (embedded) in particular environment, which is shared out among several agents;

The innovation environment that has to ensure the occurrence of an innovation process can adopt an agent-oriented view, since it has decentralized nature. The realization of an innovation depends on actors including mediators that work autonomously. Moreover, these agents need to interact, in order to either to achieve their individual objectives or to manage the dependencies that ensure from being situated in a common environment [3]. In a specific innovation environment the main actors (technology transfer offices, innovation centres and others) participate in mini-societies devoted to specific objectives. The creation, operation and dissolutions of such societies are achieved by agents acting autonomously.

The exploitation of the agent-based approach to development of an innovation environment that has to assists such a complex activity as realization of innovation has an

objective to achieve a successful construction of an innovation system. A multi-agent system can be designed, where several agents offer specialized problem solving activities and interact with each other to produce the overall, complex service to realization of an innovation strategy. Each individual agent, typically an expert or a computer-based component with problem solving capabilities can interact and co-ordinate its activity with other agents of a network in order to handle effectively the common problem – innovation realization. The agents devoted to solution of specific problems concerning innovation realization can use very different intelligent approaches, so that heterogeneous multi-agent architectures could be designed.

The objective of this paper is to present a way of designing innovation environment using agent paradigm. The next section presents requirements to a multi-agent innovation environment. The third part describes the all phases during development of a multi-agent system and shows a design model of an agent. A brief description of main agents of a decentralized innovation environment is given in the next section. Finally, it is presented the information and intelligent systems of a concrete innovation agent – Technology Transfer Office "Information and Communication Technologies for Energy Efficiency".

2. Agent-oriented view to design and implementation of innovation environments

An innovation environment can be considered from two viewpoints:

1. Viewpoint of its constructor – the engineer's viewpoint;
2. User's viewpoint – it concerns innovation system usability: the viewpoints of system's clients.

The development of this environment takes into consideration the engineer's viewpoint. In keeping conformity with it, the construction of an innovation environment is the result of the modeling done during all the stages of the development process (analysis, design and implementation) and the successful transformation of the analysis model into the design model and transforming of the design model into the implementation model. In the bottom up approach to development, the implementation has influence on the design modeling and analysis modeling.

The implementation of a design model of innovation environment depends on two factors: implementation model and implementation platform. The implementation model describes a way of realization of a design model. It depends on the existing (information and communication) technologies that ensure some implementation platform. The latter is a factor of realization of implementation models. Mainframes, mini- and personal computers, network of computers, Internet, Grid technology, Semantic Web are various implementation platforms that force the construction of implementation models in different ways. The implementation platforms like Internet, local networks, Grid and Semantic Webs give possibility for realization of the distributive variant of an innovation environment. The multi-agent technology is suitable for construction of its agent-based implementation model, since an agent is considered usually as an entity realized by encapsulated computer system that is situated in some environment [4]. The multi-agent paradigm is also useful framework for constructing the design model of an innovation environment.

The fundamental characteristics that describe this paradigm depend on the basic definition of an agent, but it is possible to reveal additional agent characteristics due to different ways of acting of an agent (agency):

1. An agent can be *reactive* or *proactive*. An intelligent agent is proactive, if its agency is the capacity of human beings (decision making entities) to make choices and to impose those choices on the world through their operation.

2. The agency is a *collective*. An agent can have a complex structure – it could consist of several other agents.

3. If the agency determines agents that do not act alone, their nature enables their description in the framework of the Actor-Network theory [5]. According to it, the agent is an actor that performs a *specific role*.

4. Agency can present *agent's operation as tasks' performance*. In this case the agent-oriented design can use task analysis technique for analysis of work domain covered by agents. There are various task analysis techniques that are classified as device independent/ dependent, event dependent/ independent and psychologically relevant/ irrelevant [6].

6. The agent has capacity to act accordingly special knowledge that it possesses, i.e. its agency enables it to solve problems during its operation. Its *problem solving* capabilities can range from agents with fixed behavioral rules to agents with problem solving ability [7, 8]. The former perform a task in correspondence with a model of behavior in a constant, predictable environment that allows some adaptation. The latter are capable of exhibiting flexible problem solving behavior in pursuit of their design objectives. They operate in an environment, over which they have partial control and can solve new undetermined problems. In this case the development of multi-agent technology and agent-oriented design modelling can take advantages of a useful descriptive framework related to Distributed Cognition paradigm [9]. This descriptive theory is a branch of cognitive science that proposes knowledge is not confined to an individual. Rather, it is distributed across individuals, and artifacts in the environment.

3. Phases in multi-agent system development

A multi-agent system is distributed and concurrent. Agents can be heterogeneous in terms of used knowledge model, their capabilities and structure. While the user of such system is not aware of its complexity structure, the developer has to come to terms with the difficulties related with the understanding of system behavior. The development of such a system passes through the following phases:

- Construction of single problem solving node (agent);
- Building the interactive features of individual agent;
- Building of integrated system.

The accomplishment of each one of these phases requires of the developer to have mental pictures of several aspects of the system. They are result of its analysis from the following viewpoints – *Agents*, *Agents Interaction*, *Agents Collaboration* and *Problem Domain Consideration*. The examination of Agents supports the presentation of single agent, its structure, behavior and problem solving activity. The result is an Agent View. The look at Agents Interaction and Collaboration provides information that is needed during the second and third phases of system developing, respectively. This information serves for presentation of Interaction View and Collaboration View on the system and innovation environment. Problem Domain Consideration results in Problem Domain View that is necessary in construction of agents and building of integrated system.

The upper views that are introduced during the analysis of innovation environment have to be represented in several developer's conceptual models, which provide the theoretical basis of the multi-agent system. It drives the determination of a number of alternative specifications of system elements. We concentrate on the following developer's conceptual models:

1. Domain (Task) Model. It presents the characteristics of the problem. Dynamically changing problem attributes attached to various agents might indicate the places where to take correct actions.

2. Agent Model. It can take different forms – a structural agent model and behavioral agent model.

3. Interaction Model. It describes the message flow among and within agents characterized by types, contents, time and etc.

4. Cooperation Model. It concentrates on the collective performance of problem solvers.

5. Model of agent competence. It consists of a knowledge and skills models. They represent what agents know about themselves and about their capacities during problem solving.

6. An Architecture Model that represents structurally individual agents and groups of agents.

The relations among the developer's conceptual models of a multi-agent system and different views that present an innovation environment in multi-agent framework can be seen on Fig. 1. As the individual agent is fundamental entity of a multi-agent system, the Agent View supports the construction of several conceptual models that could be seen as agent representation at several levels. At the highest level is the structural Agent Model that describes the main components of an agent and their relations. Since the agent is described as a problem solver the agent view has to give information about the tasks that it has to perform and the knowledge that it is needed to carry out intelligent functions. This information is useful for construction of Model of agent competence. The development of an Architectural Model is result of Agent View, as well. The Interaction View is useful for creation of the Interaction Model that bases on a number of ways of presentation of inter-agent interaction. The Collaboration View serves to modelling of agent co-operations. The Cooperation Model represents the static and dynamic task dependences among agent that are determined by Problem Domain View. The latter is source for construction of Domain Tasks Model. It serves for development of Behavioral Agent Model and description of intelligent functions of agents, since domain problem solving can be represented by a hierarchical decomposition of tasks, which are necessary for achieving the problem solution.

The presented conceptual models suggest the design model of a node of a multi-agent system, i.e. the design model of an agent. It has to consist of components that are intended to realize the developer's conceptual models in whole or partially. Figure 2 presents such a design model. According to it an agent consists of two layers – *Decision Making Layer* and *Cooperation Layer*. The Decision Making Layer has two components: an intelligent system and information system. The intelligent system performs domain problem solving tasks. Some of them concern decision making. Their execution bases on information availability. The latter is provided by an information system that stores all necessary data and knowledge get by the agent. The Cooperation Layer is presented by a Communication Module and agent's control.

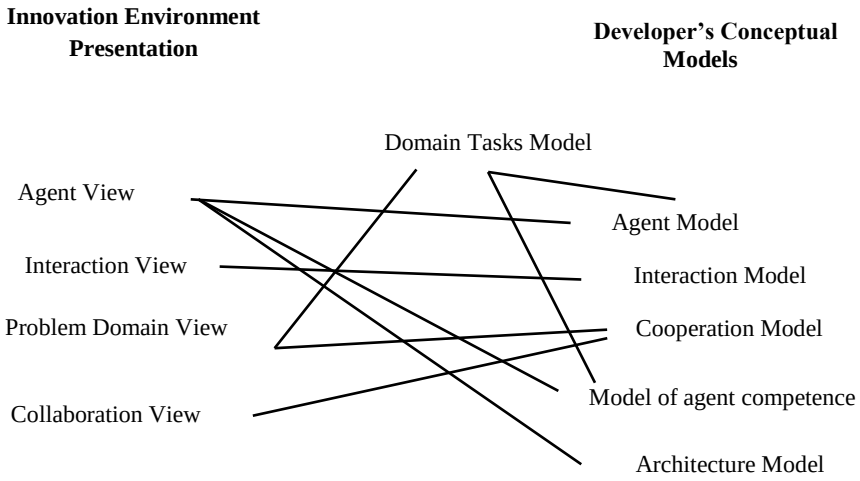


Fig. 1 Relations among environment presentation and conceptual models

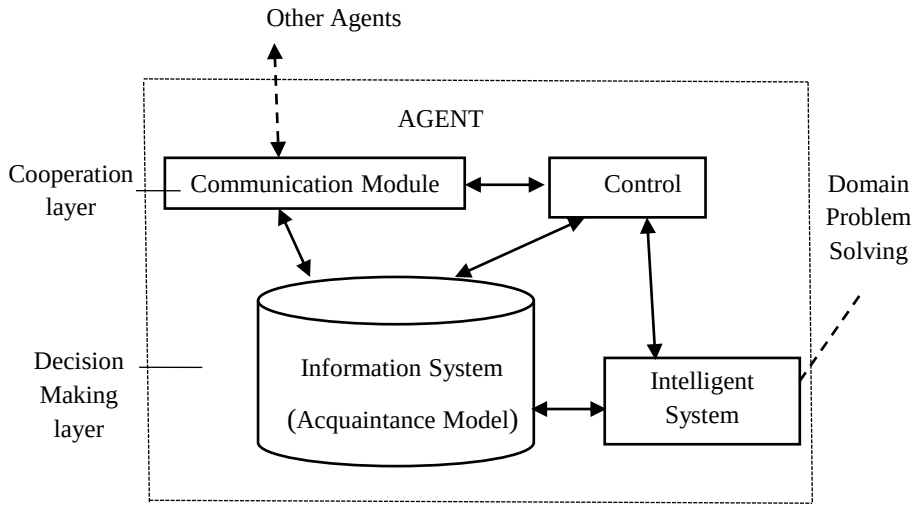


Fig. 2 Design model of an agent

The control module directs the intelligent system in such a way that the individual agent provides the necessary services, which it is responsible to ensure in accordance with its objective and role in the multi-agent system. As a part of the cooperation layer it controls the individual agent in such a way that it demonstrates co-ordinate behavior towards the other agents and contributes to the integration of its intelligent system in the multi-agent environment. In other words, the cooperation layer manages interaction with the other agents and relates the local activity of the intelligent system to the global problem solving. This

achieved through a control mechanism, which uses a set of data structures modelling the agent's acquaintances (acquaintance models). Information (knowledge) contained in the Information system is used for planning co-ordinate activity and other social interaction. The interaction with the other agents is done by the communication module, which performs intelligent filtering and massage routing.

4. Agent-oriented design model of innovation environment

Since the innovation environment has collaborative framework, it could be designed through a multi-agent system. In this way all participants in the support of an innovation process act autonomously in solving specific problems. The agents are autonomous entities for innovation problem solving. They realize modular systems that integrate heterogeneous components. The agents are usually embedded in a specific environment to carry out autonomously predefined functions concerning innovation. Their flexible action in dynamic environment is one of their characteristics. Since the collaboration is supported by different agents, the innovation environment can be described by numerous agents that have to interact with each other. There exist different relationships among them – working together as peers or building hierarchical structures with "managers" and "members". Multi-agent systems have different architecture, control strategy, interaction protocols and agent complexity.

The design of multi-agent system bases on presentation of a conceptual model of the innovation environment and elaboration of detail models of each particular agent. The conceptual model defines the multi-agent architecture/structure in accordance with an innovation environment and its purpose, agents' roles and communications. The agents' roles depend on the problems that they have to solve and the system structure. Each role is designed to fulfill particular objectives and/or to solve particular problems. That is why the agent's functioning has to be fully and clearly specified by internal states diagrams, their behavior, objectives, interactions, etc.

Table 1 shows some of the agents in an innovation environment that serves for supporting of innovation. Their main objectives, activities and interactions are outlined.

5. Technology Transfer Office "Information and Communication Technologies for Energy Efficiency" – Intelligent System and Information System

The main objectives and activities of the agent "Technology Transfer Office "ICT for Energy Efficiency" (TTO "ICTEE") determines the problems that have to be solved by the office. Their solutions require the performance of tasks that is responsibility of the intelligent system of the agent. These tasks can be divided in the following four groups of services provided by the agent:

1. Information Services:

- Development of a database of energy efficiency related micro-, small and medium-sized enterprises;
- Constant networking with other Technology Transfer Offices from the National Innovation System of Bulgaria;
- Advancement of novel technological and scientific achievements in the field of energy efficiency and implementation of ICT;

- Support in promoting companies and their products;
- Establishment of contacts with research and development organizations and enterprises (foreign ones included); assistance with the establishment of collaborations between local and foreign companies;
- Organization of workshops, seminars, exhibitions, etc.

2. Consultancy Services:
- Evaluation and assistance in the protection of intellectual property and patent rights, legal support;
- Development, consultancy and coordination of new project proposals concerning the usage of ICT in the area of energy efficiency;
- Documentation development, certification and registration of new products;
- Market analysis of the market of ICT based energy efficiency innovation products.

3. Financial Services:
- Facilitation of easier access to external funding and support;
- Access to finance schemes;
- Promotion of energy efficiency related developments towards investment funds.

4. Technological Services:
- Transfer of technologies, innovations and know-how, scientific solutions to particular problems, related to the usage of ICT for energy efficiency;
- Technology valuation and technology audits;
- Pilot testing of energy efficiency technologies;
- Experts and consultants services, related to technology transfer.

The information system provides necessary information for performance of the upper tasks. Since the role of the TTO "ICTEE" is to be a mediator among the various types of TTO's clients, it is very important that the information system has to store information about them. The agent works with two types of clients:

1. Organizations - Research organizations; Engineering companies; Factories; Government institutions; Community, non-governmental organizations.

2. Individuals - Members of professional groups; Represented individually.

They are from the following target economic sectors:

- Software development;
- Research and development of computer and communication systems;
- Engineering;
- Energy industry;
- Transportation;
- Construction and real estate.

Table 1. The main goals and activities of the agents

Agent	Objectives	Activities	Interactions
Technology Transfer Office "ICT for Energy Efficiency"	*To be a factor for the development and enhancement of the 'science to business' relation, promotion of new knowledge *Technology transfer towards Bulgarian enterprises in support of their competitiveness and sustainable development *Realization of innovation strategy in the sectors of Energy Efficiency (EE) and Information and Comm. Technologies (ICT)	*Assessing possibilities for improving energy efficiency through the use ICT in different economic sectors such as energetics, machine-building, transportation, construction, etc. *Participating in networks of TTO as well as in existing innovation infrastructures *Providing consulting and expert services in the field of technology transfer, intellectual property and financial management of innovation projects *Disseminating industry-related results of the applied research conducted by institutes of BAS countrywide/worldwide.	Other agents– mediators in innovation environment
Technology Transfer office (in Bulgarian academy of sciences have 15 TTO)	Use a model of technological entrepreneurship	*__Initialising__: The quest for new business and the procurement of resources are the key to the entrepreneurial initiatives of the transfer entrepreneur *__Combining:__ The conscious and rapid adaptation to ongoing changes as part of the entrepreneurial behavior of the transfer entrepreneur requires firstly the coordination of objectives and resources such as personal network, secondly the combination of relationships and thirdly the creation and extension or conversion of the organization. *__Implementing:__ The steadfast pursuit of objectives despite failures is an essential key to the entrepreneurial implementation of the transfer entrepreneurial. [10]	Other Steinbeis Transfer Centres
Technology transfer office that transfer scientific results of the research organization	Use a lot of intermediary function to solve a market problems	*Diversification of industry-related results by the applied research conducted by institutes of BAS countrywide/worldwide. *Dissemination of this science results	Other transfer offices and SMEs

6. Conclusion

The objective of this paper is not to describe an agent "TTO "ICTEE" in detail. It has to present a way of designing an innovation environment on the base of agent paradigm. It is a useful approach, if we want to realize a virtual innovation environment that consists of virtual agents.

7. Acknowledgments

This document is created with financial support of the Project "Construction of Technology Transfer Office "Information and Communication Technologies for Energy Efficiency", contract No: BG 161PO0003-1.2.02-0001-C000 with the Operational Programme "Development of the Competitiveness of the Bulgarian Economy" 2007-2013, financed by the European Union via the European Regional Development Fund. The entire responsibility for the contents of the document rests with the Institute of Information and Communication Technologies, BAS, and under no circumstances can it be assumed that this document reflects the official position of the European Union and the Contracting Authority.

References

1. Jennings N. R., An agent-based approach for building complex software systems, *Communications of the ACM*, vol. 44 (4), pp. 35-41, 2001.
2. Singh M. P., Conceptual modeling for multiagent systems: applying interaction-oriented programming, in *LNCS*, vol. 1565, P. P. Chen et all, Eds. Berlin Heidelberg: Springer-Verlag, 1999, pp.195-210.
3. Foster I., N. R. Jennings, C. Kesselman, Brainmeets brawn: why grid and agents need each other, in *Proceedings of AAMAS'04*, July 19-23, 2004, NY USA
4. Wooldbridge M., *Agent-based software engineering*, in IEE Proc. Software Engineering, vol. 144, 1997, pp. 26-37.
5. Latour B., Reassembling the Social: An Introduction to Actor-Network-Theory, Oxford University Press, Oxford, 2005.
6. Vicent K. J., Wanted: psychologically relevent, device- and event-independent work analysis techniques, Interacting with Computers, vol. 11, 1999, pp. 237-254.
7. Bosser T., *Models of user behaviour and measures of usability*, in NATO ASI Series F, vol. 129, Human Machine Communication for Educational Systems Design, M. D. Brouwer-Janse, Th. L. Harringtton, Eds. Springer-Verlag, 1994, pp. 89-97.
8. . Van Liedekerke M. H, N. M. Avouris, *Debugging multi-agent systems*, Information and Software Technology, vol. 37 (2), 1995, pp. 103-112.
9. Perry M., *Distributed Cognition*, In J. M. Carroll (Ed.) HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science. Morgan Kaufmann, 2003, pp. 193-223.
10. Auer Michael, *Transfer Entrepreneurship*, 2010.

Analytical Representation of Graphs by Means of Parametrically Defined Splines

Tatyana Romanenko[1], Vitaliy Vishnevskey[1],Vladimir Kalmykov[1]
[1] Institute of Mathematical Mashines and Systems Academician Glushkova Avenue 42, 03680, Kiev, Ukraine
vl.kalmykov@gmail.com

Abstract. Very often the experimental data reveal the process which is fully determined by some unknown function, have been distorted by noises. Treatment and experimental data analysis are substantially facilitated, if these data to represent as analytical expression. Usually splines are used to approximate functions that are given in tabular form. Among the various types of splines, Bezier curve is most convenient to approximate the graphs. The algorithm of experimental data processing and the example of using this algorithm for spectrographic analysis of oncologic preparations of blood is represented in this article.

Keywords: graphics, experimental data, spline, Besier's curve.

1 Introduction

The experimental data usually represent the measurement process, which has been distorted by noise. The process is fully determined by some unknown function $y= f(x)$. Let the measured value corresponds to a single-valued function $y = f(x)$, and as a result of measurement, we have a finite sequence of N pairs $\{x_i, y_i\}$; $i = 0,N$. In most cases experimental data are represented as the graphical curves. The graphical curves , i.e. graphs are, apparently, the simplest and a long ago in-use means of cognitive presentation of experimental data in the most different scopes of human activity which allow to estimate clearly the qualitative property of the process, in spite of clutters, measurement errors. Graphs, displaying the same processes or objects, can substantially differ from each other by scales, amount of the used measurements, level noises, and so on. At the same time the feature form of the graphical curve characterizes the parameters of the displayed object or process.

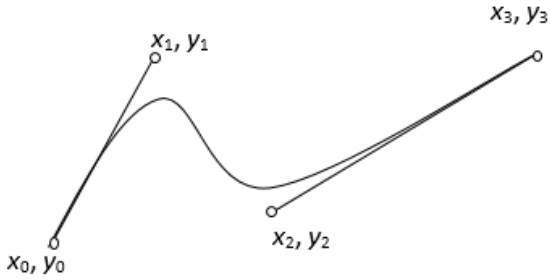


Fig. 1. Bezier curve

Automatic (automated) processing of the graphical curves supposes comparison of their forms to set,

whether the different graphs characterize the same or different processes or ojects. The direct using the neuronet methods or statistical pattern recognition methods to solve this task, is impossible, because different graphs relating to the same object can differ from each other on such parameters as scale, noise, number of measurements, etc. Thus, the graphs should be approximated by functions that are invariant to affine transformations to make it possible to compare the graphs. Usually splines are used to approximate functions that are given in tabular form [1]. Among the various types of splines, Bezier curve (Fig.1) is most convenient to approximate the graphs. So, the experimental data – graphical curves may be replaced with the analytical curves that are congruent enough to the initial graph form, and, at the same time, invariant relatively to the scale change, amount of measurements, noises level. Such curves may be chosen as the polynomial functions $x(t)$, $y(t)$ of n power of the parameter t :

$$\begin{aligned} x(t) &= a_0 t^n + a_1 t^{n-1} + \ldots + a_{n-1} t + a_n, \\ y(t) &= b_0 t^n + b_1 t^{n-1} + \ldots + b_{n-1} t + b_n. \end{aligned} \qquad (1)$$

Choice of n – the power of polynomials depends on complication of the experimental curves to be approximated.

If such approach would be realized it is possible to hold all the substantial form features of the experimental curve and to eliminate the noises influencing. In addition, instead of the graph description in the space of initial signals it is possible to operate approximating polynomials in the space of their coefficients. The substantial advantage of such description is its invariance relatively the amount of the measured signals in every graph.

A method to approximate the graphics splines is known for a given amount of points is not large, tentatively – up to 10 [2]. It is believed that the spline curve must contain all of the given graphs points. In this case, the system of N equations must be solved to obtain the polynomial coefficients. But if the number of points in the graph is a few hundred, and in the presence of noise, the problem ceases to be trivial and does not have the simple decision [3]. One of this task decisions is proposed in the article.

2. Format Problem Definition

The graph can be represented by a polyline, the kinks of which form the sequence of N experimental values of some function $y(x)$. Let us consider the Bezier curve as an approximating curve because of the most often using in practical applications, parametrically defined polynomial of the third degree.

Under the Bezier curve we understand the arc of flat curve of the third degree [3]:

$$Bz(t) = \sum_{j=0}^{3} Br_j(t)\, Q_j \qquad (2)$$

where $Br_j(t)$ are base Bernstein polynomials of the third degree,

Q_j are the coefficients of curve t is an real parameter which changes in an interval [0,1].

Bernstein polynomials are defined as:

$$Br_j(t) = C_3^j (1-t)^{3-j} t^j, \quad 0 \le j \le 3 \qquad (3)$$

where $C_3^j,\ 0 \le j \le 3$ are the binomial coefficients

$$C_3^j = \frac{3!}{j!(3-j)!} \tag{4}$$

Thus, the Bezier curve (fig. 1) can be set by two polynomials of the third degree depending on the parameter t:

$$\begin{aligned} X(t) &= x_0(1-t)^3 + x_1(1-t)^2 t + x_2(1-t)t^2 + x_3 t^3 \\ Y(t) &= y_0(1-t)^3 + y_1(1-t)^2 t + y_2(1-t)t^2 + y_3 t^3 \end{aligned} \tag{5}$$

The experimental data approximating as Bezier curve has many of advantages. Describing the wide enough class of curves, Bezier curves use reasonable quantity of coefficients are 8. Coefficients (x_0,y_0), (x_3,y_3) are the co-ordinates of initial and eventual points of approximating curve arc, accordingly, and coefficients (x_1,y_1), (x_2,y_2) are the co-ordinates of points which control its form (so-called "whiskers"). Changing coefficients or moving points proper to them on the screen of monitor (both initial and eventual and controls), an user can edit the form of approximating curve in the interactive mode, improving the results of automatic approximation.

Thus, every graph may be represented as a vector $v = \{ x_0,y_0, x_1,y_1, x_2,y_2, x_3,y_3\}$, and its components are the coefficients of Bezier curve equations. The graph description as the vector v, always has the same amount of components regardless of number of measured experimental points, that makes solving of different tasks, in particular, tasks of the graphs recognition more simple.

It was noticed before, that determination of Bezier curve coefficients, most exactly approximating some graph, is not a trivial task. The iterative method is used to determine the coefficients of Bezier curve, approximating some experimental sequence of points, by minimizing the distances squares sum from every point of the graph up to the approximating curve[3]. Such a formulation of the problem follows from the assumption that the considered graphs are formed by mutually independent points, and fact, that these points describe the object as a whole, is not taken into account. To calculate the distance from every point of the graph to the approximating curve, proper nearest point on the approximating curve must be found. Also, this task is iterative as the approximating curve is parametrically defined. When the quantity of points on the graph is large (more than 100), and the amount of iterations is not limited beforehand, such a method of task solving has resulted in substantial calculable difficulties. Also, the spline approximation in [2] uses the assumption of mutual independence of the experimental data points, but the computational problems do not arise when using the method due to the small number of points.

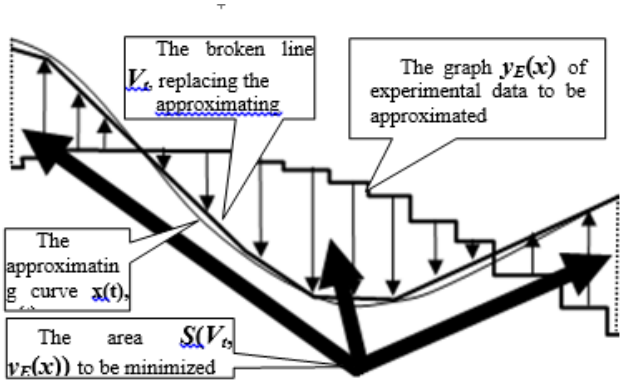


Fig. 2. The area to be minimized between the broken line V_t, replacing the approximating curve $x(t)$, $y(t)$, and the graph $y_E(x)$ of experimental data to be approximated.

On the contrary we believe the right to consider all points of the experimental data related because they belong to some function. Consequently, the neighboring points of the graph must be connected by line segments. Thus, the graph of experimental data is a broken line.

We propose to accept the area $S(v,y(x))$of the figure, limited by the contour, formed with the approximating and experimental curves as the measure of likeness of this experimental graph with the approximating curve. So the values of Bezier curve coefficients v_{opt}, which approximates certain experimental graph in the best way, correspond to the minimum value $S(v,y_E(x))$:

$$v_{opt} = \arg\min_{v} S(v, y_E(x)) \tag{6}$$

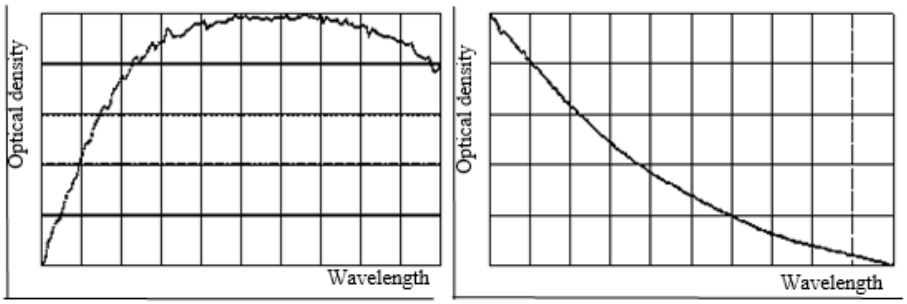


Fig.3. Examples of blood preparation spectrograms: on the left is spectrogram of sick, on the right - healthy man.

3. Algorithm

It is follows from (6) that computation of the area which is restricted by the contour $S(\boldsymbol{v},y_E(x))$, is basic operation of algorithm for searching the optimal coefficient of $\boldsymbol{v}_{opt}$. The direct calculation $S(\boldsymbol{v},y_E(x))$ is impossible, since this contour is not simply connected.

The graph $y_E(x)$ and Bezier curve intersect each other many times. It is difficult to compute the co-ordinates of intersections because of Bezier curve is parametrically defined.

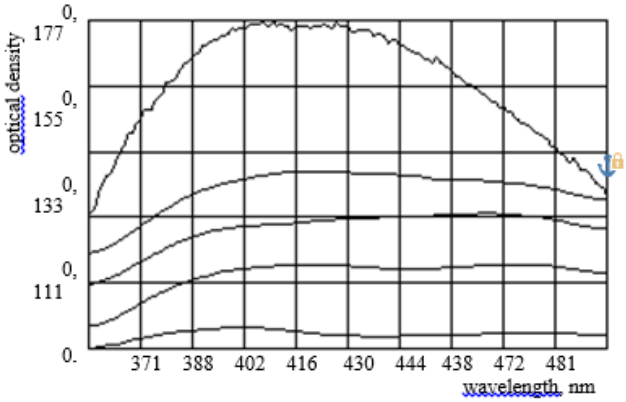


Fig. 4. The blood preparation spectrograms if the oncology diseases are of the same localization (stomach).

We propose to replace the Bezier curve to a polyline V_t. The fracture points of the polyline (x_t,y_t) belong to that Bezier curve and are calculated for the $t = 0, \delta, 2\delta, 3\delta,...,1$. The value δ is chosen sufficiently small, in order while replacing of the initial curve to the polyline Vt, the accuracy to be enough for the practical calculations. As a result, made to simplify the explicit value of the approximating polyline $V_t(x_n)$ directly corresponds to each value $y(x_n)$ of the graph (Fig.2). Then the area $S(\boldsymbol{v}, y_E(x))$ can be calculated as $S(\boldsymbol{V_t}, y_E(x))$:

$$S(V_t, y_E(x)) = \sum_{n=0}^{N} \left| V_t(x_n) - y_E(x_n) \right| \quad (7)$$

The value $\boldsymbol{v}_{\mathbf{opt}}$, corresponds to the minimum value $S(\boldsymbol{v},y(x))$, can be got using the gradient method.

4. Experimental Verification of Algorithm

The proposed algorithm was experimentally tested on more than 1,100 real the graphs of the spectrograms, and preparations were obtained by the method of Oncotest [4]. The method to diagnose early oncology diseases allows to define presence or absence of oncology disease using the spectrograms of blood preparations (Fig.3). At present researches are conducted also and for determination of localization of disease. The spectrograms of blood preparations are represented in Fig.4. Their diagnoses are the oncology disease of identical localization (stomach). In spite of obvious distinctions of spectrograms on the levels of signals, hindrances, form of curves have undoubted likeness. The form of curves is specific property of the resulted graphs, which can correspond to the diagnosis. The identification task to form a diagnosis by spectrograms supposes additional researches which must be based on treatment of plenty of experimental information. In the same queue, the processing of large data arrays is impossible without the decision of tasks of their computer treatment, and including the tasks of automatic or automated classification of spectrograms on their form must be decided, using the methods of patterns recognition.

The direct use of statistical pattern recognition methods for the automated diagnosis forming in this case is not possible owing to large changeability of scale (dozens of one times), and, also, configuration of function of realizing the same appearance. Therefore it is appeared expedient to find some suitable Bezier curve for every graph, which polynomial coefficients, after adduction to one scale, can be used as pattern parameters, invariant to the changes of scale, amount of measures, noises level.

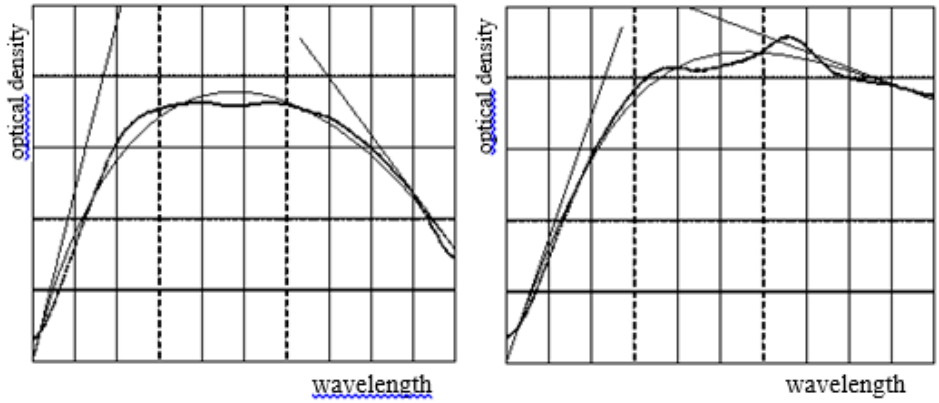


Figure 5. Spectrograms of blood preparations and their approximation by Bezier curves; localization is a mammary gland

Examples of spectrograms and approximating them Bezier curves are represented on a fig. 5, 6. A result was achieved in most cases for 15 - 40 iterations (in especially difficult cases - to 300). As it is obvious from figures, approximating the Bezier curves give the satisfactory approaching.

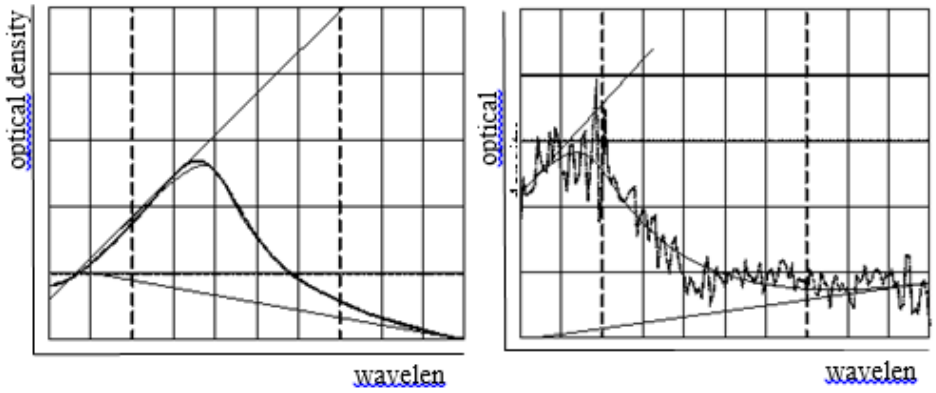


Fig. 6. The spectrograms of blood preparations and their approximation by the Bezier curves; localization is a liver

5. Conclusion

A new method for approximating the graphs that reflect experimental data is proposed. The proposed method differs from the known methods in that it is based on the assumption that points of the graph represent a function as a single whole, and therefore are connected with each other and therefore a connected path is formed.

There is proposed to use the size of the area between the approximating function and the graph in order to evaluate the their closeness.

To simplify the calculations, we propose to replace the approximating function to broken line, the nodes are the points of this function, and the nodes are chosen so that the accuracy of the approximation corresponds to specified requirements.

The offered algorithm allows finding the parameters of Bezier curve, which approximates the experimental graphs.

The examples of approximation of spectrograms are resulted for implementation of the automated diagnostics. It will enable in the future to automatize the process of screening examination of population with the purpose of early detection of malignant tumors.

References

1. J.H. Ahlberg, E.N. Nilson and J.L. Wash, The Theory of Splines and Their Applications, Academic Press, New York, 1967
2. D.F. Rogers, J.A. Adams, Mathematical Elements for Computer Graphics, McGraw- Hill, New York, 1976
3. V.V. Vishnevsky, I.K. Rystsov, and M.V. "Voljeva, Iteratsionny algoritm postroyeniya krivoy Bezie po zadannymi tochkami", Matematicheskie mashiny i sistemy, no. 4, pp. 108-116, 2004 (in Russian).
4. V.V. Vishnevsky, V.A. Vladimirov and T.N. Romanenko, "Programmno-apparatny kompleks Oncotest –WM-01", Ukrainian Journal of Telemedicine and Medical Telematics, vol. 4, no.2 pp. 182-185. 2006 (in Russian).

Fostering Innovation in Cultural Industry: The Case of Bulgarian Capital – Sofia

Tsvetelina Yorgova[1], Kostadin Kostadinov[2]

[1] BAS, 1113 Sofia, Bulgaria
tsveti.yorgova@gis-tc.org

[2] IMech - BAS, Acad. G. Bonchev Str., Bl. 4 1113 Sofia, Bulgaria
kostadinov@imbm.bas.bg

Abstract. This paper analyses the role of technology transfer offices (TTO) in the innovation process in cultural industry and propose how to promote innovations in culture. The Innovation in cultural industries is generally considered as non-technological one, which takes higher and higher share of GDP in the EU. We consider that the cultural industry is all around us; it is everything that forms in a certain way our values, scopes our moral principles, and gives us knowledge of history. TTO takes place in the innovation system not only as an intermediary in the communication between users of cultural innovation and authors, but as a promoter of entrepreneurs in culture to innovate. In this case, the TTO is a tool to provoke innovative thinking among the professionals in the field of cultural heritage and also provoke creative thinking among the young people, students and others. In this article we suggest a way to make innovation to client (in our case-Sofia municipality) without preliminary defined market needs or demands. To promote the creativeness for cultural innovation and to improve the competitiveness of Sofia's candidature while preparing for competition European Capital of Culture 2019, a proactive approach has been used, which is based on a competition for cultural innovations in desired fields. These fields of culture industry are defined by a group of experts in specific topics of cultural industries. The method of open competition has been selected for implementation to attract the society and to encourage the active and creative work of the potential participants. In conclusion recommendations are made what a local cultural industry strategy might be, as being used the developed methodology for facilitating broad spectrum of innovators.

Keywords: Cultural industry, non-technological innovation, technology transfer office

1. Introduction

Cultural Industries are defined by UNESCO (Unesco, 1982) as those industries which produce tangible or intangible artistic and creative outputs, and which have a potential for wealth creation and income generation through the exploitation of cultural assets and production of knowledge-based goods and services (both traditional and contemporary). All cultural industries have in common that they all use creativity, cultural knowledge, and intellectual property to produce products and services with social and cultural meaning.

The cultural industries may include: advertising; architecture; crafts; designer furniture; fashion clothing; film, video and other audiovisual production; graphic design; educational and leisure software; live and recorded music; performing arts and entertainment; television, radio and internet broadcasting; visual arts and antiques; and writing and publishing. The term "cultural industries" is almost interchangeable with the concept of "creative industries." Whereas the notion of "cultural industries" emphasizes those industries whose inspiration derives from heritage, traditional knowledge, and the artistic elements of creativity, the notion of "creative industries" places emphasis on the individual and his or her creativity, innovation, skill and talent in the exploitation of intellectual property. The notion of 'cultural industries' is also closely linked to but, again, slightly different from a categorization based strictly on the notion of "intellectual property," which is closely linked to the concept of information-driven or knowledge-based economies, and which includes such activities as scientific and technological innovation, software and database development, telecommunication services, and the production of hardware and electronic equipment.

Cultural industries are also deeply linked with the innovation (Ragot, 2006) generally in two phases. During the first phase the cultural product or service is created or produced. The second phase is the one that keeps the cultural industries closer to other industries. In this phase the technological innovation can create a new process of production, as well as a new service support for the creative content. The first phase is non-technological innovation, which nowadays takes higher and higher share of GDP in the EU.

The public authorities play significant role in the cultural industry. There are five ways in which the public authorities can achieve the major objectives concerning cultural industries (Unesco 1982):

- Direct aid, consisting mainly of subsidies or purchases of products and services out of the state or municipality budget;
- Indirect aid in the form of tax relief, varying a great deal from one country or municipality to another;
- Regulations (specifications, programme contracts, etc.) which compel certain branches to comply with rules governing public service, such as for example the maximum percentage of foreign films or songs and so on;
- International and national conventions, especially concerning copyright (authors and artists);
- Other incentives like festivals, prizes, lotteries, etc.

Here in this paper we will apply another one approach using the technology transfer for fostering innovation in cultural industry for the case of Bulgarian capital- Sofia.

2. Problem definition

GIS-Transfer center seeks to redefine the traditional technology transfer process between the science and industry and vice versa (Popov, Kostadinov 2005) by creating an unconventional entrepreneurial solution. The key to success is the establishment of an open system in which competent partners with diverse background combine their complementary expertise to create technological innovations that lead to successful venture enterprise provoked by the related technology transfer.

The problem to promote innovations in cultural industry arises from the necessity to elaborate competitive application of Capital city of Sofia for the competition of Cultural

Capital of Europe 2019. For this purpose a project managed by prof. K. Kostadinov in GIS-Transfer Center on developing and improving non- technological innovations has been applied to program Europe of Sofia Municipality for funded and has been further implemented.

The main objective of this project is to encourage the public to generate innovation for a dynamic cultural industry of the city with an end result of increased competitiveness Sofia's bid for European Capital of Culture and increase its potential in economic and cultural aspect. Target groups are:

- Citizens of the capital city, which will be encouraged and assisted in the development of innovation in the cultural industries;
- Regional administrations, which will be supported and activated to serve citizens, not only for innovation in the cultural industry, but also for their involvement in initiatives and programs in the application of Sofia for European Capital of Culture.

According to Priority area 2 of the program Europe the project contributes to public engagement to generate innovations for a dynamic cultural industry in Sofia, and for the competitiveness of the candidature of Sofia for European Capital of culture. To support this it is necessary to define such priorities, which can in best way discover the potential of the city and to present it to the European citizen in 2019. The leading experts in culture, economy and innovation have been selected to define the most appropriate priorities to fulfill the project aims.

3. Approach used in project implementation

We use the incentive or promotional approach for technology transfer to encourage the creativeness of citizens. Here we describe how it is realized.

- Selection of current topic of the cultural heritage of the city, which can be implemented innovative approaches and experimental development. These are explored as well as established and emerging sector and specific sites of cultural heritage and the experts made choices of current and attractive ones. Selected objects evaluated for their value and originality or attractiveness, including and tourist interest. The projects are: buildings- architectural monuments containing historical value; churches and prayer temples with their history, such as the tetrahedron of religious tolerance (Orthodox church 'St. Nedelya"- Mosque - Synagogue- The Catholic Church 'St. Joseph"); Museums - with national and local significance, including local crafts; Parks - North park, West Park and other local and green areas, including with opportunities for new types of entertainment and relaxation; Touristic or hiking trails;
- Tools for fostering: Announcing the public competition for innovative project proposals, approaches and tools for their promotion in our society and among the tourists and guests of the capital. We draw the project proposals through announcing the competition through the media, incl. with prizes for the winners. For this competition we have prepared and implemented methodology for evaluating of proposals, which is made by competent jury, consisted of experts from the fields of culture and innovation.
- To promote the results of the competition a printed booklet, presenting the collected innovative ideas and projects, as well as guidelines and activities for the diversification of the cultural tourism industry in the capital has been disseminated;

- Some of the innovative proposals are selected for the participation in the exhibition and the final conference;
- The exhibition with a final conference. At the end of the competition an exhibition and conference are organized, which have been visited by a wide circle of media, citizens, contractors and participants. Except the announced awards by GIS-TransferCenter, the Association for development of Sofia and regional authorities also have awarded some of the competitive projects with the opportunity to implement their innovative ideas and projects in practice. The awarded proposals are implemented experimentally in the cultural, incl. tourism industry of the city.

4. Results and discussion

a. Convergence across industries

The cultural industry will further integrate with other industries. Convergence across the cultural, manufacturing and service industries will spur the development of all these industries. Substantial progress will be made in the integration of telecommunications networks, information and cable TV networks, and the Internet.

One example of innovative projects is Arte classica - the project is to promote classical music among the wide public through innovative ideas and practices that enhance the access of the suburbs of the capital Sofia to the classical music.

Figure 1. The Mayor of Sofia Mrs. J. Fandakova at the exhibition while discussing one of the awarded projects "Sofia treasury" with its author architect D. Mehandjiev and project coordinator Mrs. Ts. Yorgova.

b. A 'digital boom'

Digital industry will develop at fast pace and ultimately become the core of the cultural sector. Some innovative projects in this topic have been successful evaluated.

Digital technologies continue to transform our approach to creativity and culture. They modify the way cultural assets evolve and preserve. They also provide us with new tools to create cultural innovations better and faster. For example, development of an interactive map of cultural and historical assets can be compatible with all internet browsers, mobile devices and tablets. This map can combine with the integration in a centralized multilingual website. Integration of this whole set of information with geographical coordinates on a map of public transport and a map of the region with an opportunity for GPS functionality is a new information tool with more cultural and information features. Such application for smart phones can be used as an opportunity for a virtual walk through these sites. Most of the cultural assets in Bulgaria haven't been adapted for visits of disabled.

Figure2. N. Popovski - Author of one of the awarded projects "Great chess queen" with the project coordinator Ts. Yorgova

Figure3. Prof. V. Sainov and Prof. S. Sainov presented their project "Holographs copies of archeological and cultural objects" protected by several patents.

c. outcomes

37 project applications have been registered for the open competition in the framework of the funded project by Program Europe of the Municipality of Sofia.

The developed working materials - forms, criteria and methodology, which formalizes and supports selection, and evaluation of innovative ideas and projects are proposed to the regional municipal administrations and tested with their representatives.

27 project proposals have been selected to be presented at the final exhibition by a reputable jury, which has awarded three of the most innovative and with the biggest impact to the cultural industry of Sofia.

The jury has also evaluated the most innovative project ideas and proposals, which are recommended to the Development Agency of Sofia Municipality for preparation of the application for the European Capital of Culture 2019 Competition.

5. Conclusion

This paper presents the results of a funded project aiming to create not only innovations in cultural industry of Bulgarian capital city of Sofia, but to establish new sectors in that cultural industry. Innovations in cultural industry are a new area of innovation - so called non-technological innovation. Those innovations create a fundamental knowledge and values not only about the cultural-historical heritage, but for expressing and exploiting the wealth of our national identity. This is one of the challenges to Bulgarian policy (Kostadinov 2010) either national or local, facing at not only the economical, but the cultural and national identity problems, too.

Cultural industries are highly competitive, yet all countries have their own particular competence and experience based on each country's history, culture, and environment. Furthermore, through collective action and cooperation, countries can realize significant regional competitive advantages to create global economic impacts. Collaboration in regional

"creative clusters" is a strategic approach that will enable synergies between such countries and cities to promote the creative sector as a driver for development, able to penetrate even the most entrenched pockets of poverty. At the same time, a clustering strategy can serve to leverage the competitive edge of each distinct regional product in the global marketplace.

6. Acknowledgments

This work was partially supported by the Project Grant BG051PO001-3.3.06/53 of the European Social Fund and Bulgarian Ministry of Education and Science under the contract No. DO 01 -4314/13.08.13.

References

1. Kostadinov K, *Challenges for Bulgarian policy about science and innovation*, Bulgaria, 2010.
2. Popov Al., K. Kostadinov, *Technology transfer of research products*, Bulgaria, 2005.
3. Ragot S., *Innovation in the Field of Cultural Industries*, Austria, 2006.
4. Unesco, *Cultural industries - A challenge for the future of culture*, France, 1982.

Collaboration Importance in Agile Software Development

Veselin Georgiev

UNWE Sofia
vgeorgiev@petrovkata.com

Abstract. Agile methodologies in software development projects are becoming more and more popular with the development of small and medium-sized projects. This is due to several factors, which are increasingly common in modern business conditions. Different location distribution of the team members involved in the development process, the need for flexible changes in assignment due to changing the requirements, the requirement of the project owner to participate directly in the process, as well as other requirements that involve quick and flexible response from the team management and development . All these processes require many tools for coherent interaction and communication between the project participants. Collaboration tools and their capabilities evolve with the evolution of technologies that are developed and their significance for the successful implementation of the product. There are different types of approaches to collaboration, according to the type which is performed. They can be characterized relative to the manner of interaction in human resource management, in relation to the communication channels between the project participants, the need for coverage of project participants and bystanders and others.

Keywords: Agile development, collaboration, project management tools.

Agile methodologies in software development projects are becoming more and more popular with the development of small and medium-sized projects. This is due to several factors, which are increasingly common in modern business conditions. Different location distribution of the team members involved in the development process, the need for flexible changes in assignment due to changing the requirements, the requirement of the project owner to participate directly in the process, as well as other requirements that involve quick and flexible response from the team management and development. Why Agile development methodologies responds to these needs? Prof. James Highsmith called the whole process of agile software development "Agile Software Development Ecosystem", he does not like "methodology" term, because it immediately compares to traditional software development. The "Agile" process gives not only fewer processes, less ceremony, and briefer documentation, but it has a much broader perspective, which is the primary reason for using the word ecosystem rather than methodology. Although fewer processes and less formality might lower development costs, they are not enough to produce agility. Focusing on people and their interactions and giving individuals the power to make quick decisions and to self-adapt their own processes are key to Agile ecosystem and Agile development.

In internet there are a lot of interpretations for Agile process and different methods which can be defined as agile. The definition given by Wikipedia gives good overview for the whole process and idea: "Agile software development is a group of software development methods based on iterative and incremental development, where requirements and solutions evolve through collaboration between self-organizing, cross-functional teams. It promotes adaptive planning, evolutionary development and delivery, a time-boxed iterative approach, and encourages rapid and flexible response to change. It is a conceptual

framework that promotes foreseen tight interactions throughout the development cycle.". In fact Agile methodology includes different methods which can be used for incremental development, they can be very different for every single project, but that's the beauty of the ability to give power to the team participants to make decisions and changes to the development plan and product functionalities.

Agile process is best applicable for flexible projects or flexible client, whose changes project scope or requirements rapidly. In that kind of projects the main needs are periodical incremental result to be shown on the client and to be discussed next stage targets.

All these processes require many tools for coherent interaction and communication between the project participants. Collaboration tools and their capabilities evolve with the evolution of technologies that are developed and their significance for the successful implementation of the product. There are different types of approaches to collaboration, according to the type which is performed. They can be characterized relative to the manner of interaction in human resource management, in relation to the communication channels between the project participants, the need for coverage of project participants and bystanders and others.

We can define four main collaboration types between project participants during development cycle:

- **Project planning** – throughout the whole process of agile development are held planning meetings, as in the case of distributed development teams they can even be conducted virtually. In this process it is important right tools to be used so that the information to be generated may be used in the process of development later
- **Project management/task tracking** – main activity of the project managers through the development process is to track the progress of the tasks and timeframe of the project so they can take timely and appropriate action, if necessary.
- **Team inner collaboration -** within the team in addition to team members usual work communication is often done personal or not business related communication.
- **Development collaboration** - development process is most dynamic collaborative processes, main channels of communication and collaboration are between developers, developers and managers, quality engineers and developers.

To meet the needs of these 4 types of communication can be implemented the following types of tools for collaboration in software teams:

- **Software Configuration management**. A software configuration management (SCM) tool includes the ability to manage change in a controlled manner, by checking components in and out of a repository, and the evolution of software products, by storing multiple versions of components, and producing specified versions on command. SCM tools also provide a good way to share software artifacts with other team members in a controlled manner. Rather than just using a directory to exchange files with other people, with an SCM tool developers can make sure that interdependent files are changed together and control who is allowed to make changes. Further, SCM tools make it possible to save messages about what changed and why. Open-source SCM tools, such as Subversion and its predecessor CVS, have become indispensable tools for coordinating the interaction of distributed developers.
- **Bug and change tracking**. This function is centered around a database, accessible by all team members through a web-based interface. Other than an identifier and a description, a recorded bug includes information about who found it, the steps to reproduce it, who has been assigned on it, which releases the bug exists in and it has

been fixed in. Bug tracking systems also define a life cycle for bugs to help team members to track the resolution of defects. Trackers are a generalization of bug tracking systems to include the management of other issues such as feature requests, support requests, or patches.

- **Build and release management**. It allows projects to create and schedule workflows that execute build scripts, compile binaries, invoke test frameworks, deploy to production systems and send email notifications to developers. The larger the project, the greater the need for automating the build and release function. Build and release management tools can also provide a web-based dashboard to view the status of current and past builds
- **Product and process modeling**. This function encompasses the core features of what was called Computer Aided Software Engineering (CASE), from requirements management to visual modeling of both software artifacts and customized software processes. Collaboration in software development tends to be around the creation of formal or semiformal software artifacts. According to, model-based collaboration is what distinguishes software engineering collaboration from more general collaboration activities which lack the focus on using the models to create shared meanings.
- **Knowledge center**. This function is mostly document-driven and web-enabled, and allows team members to share explicit knowledge across a work unit. A knowledge center includes technical references, standards, frequently asked questions (FAQs) and best practices. Recently wiki software for collaborative web publishing has emerged as a practical and economical option to consider for creating and maintaining group documentation. Wikis are particularly valuable in distributed projects as global teams may use them to organize, track, and publish their work. Knowledge centers may also include sophisticated knowledge management activities to acquire tacit knowledge in explicit forms, such as expert identification and skills management
- **Communication tools**. Software engineers have adopted a wide range of mainstream communication technologies for project use in addition or replacement of communicating face-to-face by speech. Asynchronous communication tools include email, mailing lists, newsgroups, web forums and, more recently blogs; synchronous tools include the classic telephone and conference calls, chat, instant messaging, voice over IP, and videoconferencing. Currently, chat and instant messaging are following a similar evolution path as email communication. At first mostly used by young people for exchanging 'social' messages, these synchronous tools have been recently spreading more and more in the workplace. While email is socially blind, these tools, in contrast, provide a lightweight means to. [3]

All these approaches for collaboration in the software development project are great for facilitating normal workflows needs, but there are open problems for future usage of all generated data.

- **Different sources** – Using multiple communication systems makes our data non-centralized, which in turn makes them difficult to handle and use a shared resource analyzes in the future.
- **Not structured data** - as different sources unstructured data is also a problem while trying to process, analyze and display knowledge. Most communication tools uses plain text for distribution of the information. Plain text needs additional transformation to be done before analytic instruments.

- **Collaboration analytical tools missing** - Most of the existing tools for collaboration do not provide the tools of content analysis. Usually provide the ability to search the text, but not automatic tagging of information categorization, retrieval of knowledge or other type of text or data mining.
- **Collaborations are not used in other projects** - Typically, the completion of a software product ends usage of the information generated by the user collaboration during the development process too. It would be useful by means of analysis to facilitate the process of reusing the information where it can be added additional value for future projects.

There are some conclusions about collaboration process during software development:

- Good team and development collaboration is key requirement for successful Agile development process
- Current available project management and collaboration tools does not provide all needed functionality
- Social Network approach gives options for collaboration models development and their integration with project management and task tracking solutions
- Business Intelligent systems make it possible to build applications for analysis of information generated by the collaboration flow and to extract knowledge and added value for organization and participants.

References

1. J. Highsmith, Agile Software Development Ecosystems
2. http://en.wikipedia.org/wiki/Agile_software_development
3. Filippo Lanubile, Collaboration in Distributed Software Development, University of Bari, Italy
4. Parastoo Mohagheghi, "Global Software Development: Issues, Solutions, Challenges", University of Science and Technology (NTNU) Trondheim, Norway
5. Blair McKolskey, The problem with collaboration (http://architecturenow.co.nz/articles/the-problem-with-collaboration/#img=1)
6. Rodrigo G. C. Rocha, Collaboration Models in Distributed Software Development: a Systematic Review, Federal Rural University of Pernambuco, Garanhuns – PE – Brazil.

Information Security Risk – Analysis and Approaches

Iskren Tairov
D. A. Tsenov Academy of Economics – Svishtov, Bulgaria
itairov@uni-svishtov.bg

Absract. The present paper analyzes information systems security risk. Two types of risk analysis, their advantages and disadvantages as well as methods for combining them are described in detail. The paper advocates the idea that risk analysis is used to find a smart way for eliminating the risk or keeping it within limits that ensure the smooth flow of information processes.

Keywords. risk analysis, threat, vulnerability, countermeasures, risk management.

1. Introduction

Risk assessment is an integral part of the process of providing information security. One cannot create a security policy that reflects the specific requirements and reduces the risk to an acceptable level until a clear picture is obtained as to what the risks threatening the company's assets are. When carried out correctly, risk assessment gives the correct answer (in money and effort) regarding the results that will be achieved in security policy implementation. Building a security policy is a multi-stage process. The first thing to identify is what must be protected and then – who to protect the data and risks involved from.

2. Basics of Information Security Risk

Understanding the risk requires that several terms are clarified. The first one is the risk itself. Risk is defined as the probability of a **threat**, resulting from a **vulnerability** of the network. **Vulnerability** is the weakness of the network which may be used by malicious persons. **Threat** is any threat to the assets of the organization. When the threat has the chance to exploit the vulnerability, then there is a risk that needs to be paid attention to.

We can sum up:

Threat + Vulnerability = Risk [1]

When a company suffers losses due to a threat, this is known as **risk realization** (the term **exposure** is also used). Risk realization is what most companies try to avoid having become aware of the combination of threat and vulnerability.

Some **countermeasures** have to be taken in order to reduce the possibility of risk realization. Threats, on the other hand are potential hazards of any kind and can never be completely eliminated, one should always bear them in mind.

It can be concluded that vulnerabilities are a factor which the owners of the network have control on. Threats are such external elements that cannot be controlled. Threats should be analyzed and their impact should be assessed.

3. Approaches to risk assessment

Risk analysis can be categorized into two main groups[2]:

- **Quantitative analysis**, which attempts to attribute a value to each asset and to compare this value with the cost of the asset threat.
- **Qualitative analysis**, which requires that people have profound knowledge of the company.

Both approaches have their advantages and disadvantages. A compromise between them is ultimately reached, and this compromise is seen as the basis of the security policy.

Risk analysis[3, 4] of the information system or network security is a process that establishes the threats and vulnerabilities of an information system or network. Thus, which risks the organization is worth protecting from and which not are identified.

The process of **risk management** can begin after the analysis. Risk management is the process of reducing the risk to an acceptable level, transferring the risk to other entities such as insurance companies, averting the risk or just accepting it.

A high-level process of risk analysis consists of three steps:

- Assigning a value to assets;
- Assigning a value to risks;
- Selecting countermeasures equivalent to the value of assets and the costs of risk realization.

Assigning a value to assets

When setting values there are different elements. Some are quite obvious, while others are a little more complicated, related to external environment.

Assets assessment information can come from several sources. Finance department is one of them. There you can learn the values of the costs for acquisition, maintenance and replacement of the hardware and software. The IT staff can also be very helpful when determining the costs of disrupting the network operation and the damages that each threat may cause to the system.

Assigning a value to risks

The term **single loss expectancy (SLE)** [5] is used in the process of risk assessment when assigning a value to a particular risk. How much an asset will lose in any risk realization is determined on the basis of comparisons. When determining the value of a risk, it is not enough to define the single loss expectancy in order to make financial or resource decisions. It is also necessary to know how often the risk will be realized, or at least likely to do so.

Countermeasures

If we compare the average annual cost of any risk with the value of a given asset we can define a value that will be useful to the company in determining what resources are needed every year to protect assets against various risks. And although the purpose of risk analysis is to find an intelligent way to reduce it by taking countermeasures, in some cases, the most intelligent thing you can do is to transfer responsibility to somebody else or simply to accept this risk.

When the costs of countermeasures are too high to be borne by the company, it may decide to transfer the risk to another party. This is called getting the insurance. If the cost of restoring the network in case of a disastrous event is too large, the risk can be transferred to an insurance company which is to repair the damage after the occurrence of the event.

Should a company decide to accept the risk[6], it will simply not do anything to control the risk and will bear the costs when the event occurs.

The thorough risk analysis pursues two objectives:

- It allows security officers and financiers to make correct and informed decisions on issues affecting network security;
- It can be used as evidence of reasonable measures taken by the management to protect company resources.

In order to determine how much the costs of information security will be, a company must:

- assess the value of information assets;
- list the risks faced by these assets and the price that must be paid when these risks are realized;
- select countermeasures which are determined on the basis of their effectiveness and the costs of their implementation, compared to the cost of risk.

Quantitative Risk Analysis

The first step in risk analysis is data collection. The process of data collection should be performed methodically. First a list of all the assets of the company is made. Each asset is assigned a monetary value, which includes the cost to replace the asset, the asset's value for competitors, the impact that the asset has on the profitability of the company, etc.

The risks to which an asset is exposed must be listed. Each of these risks is related to potential losses that may be incurred if the particular risk is realized. It is very difficult to determine these losses.

Every risk is associated with the term Single Loss Expectancy (SLE). The term Exposure Factor (EF)[7] is introduced to clarify the SLE. This is a subjective, potential loss of an asset, which is expressed as a percentage and is a consequence of the realization of a particular threat. The Exposure Factor is a subjective value that the person assessing the risk must determine.

The following formula is used for determining the SLE:

Value of the asset x EF = SLE

When users, financiers and IT specialists participate in the process of risk assessment, then it is significantly easier to determine the SLE on the grounds of the funds necessary to restore the system, rather than by determining the EF.

The main benefit of using EF is that it is always less than unity, and can serve as a corrective when unrealistic rates of recovery are adopted, i.e. with EF = 1, the SLE of the risk will be the total cost of the asset.

After the SLE is calculated, the threat can be analyzed. What must be determined is how often the company is expected on an annual basis to be subject to a specific threat, provided that countermeasures are not taken. This value is known as the annualized rate of occurrence (ARO)[8]. A value of 0.0 means that during the year that threat will never be realized. On the other hand, a value of 1.0 means that it is certain that during the year the particular threat will exploit the vulnerability once. ARO = 0.1 shows that the threat exploits the vulnerability once in 10 years, whereas in ARO = 0.25 – once every four years.

The ARO is calculated in a similar way as well as the single loss expectancy – through conversations with various experts and analysis.

The **annualized loss expectancy (ALE)** can be calculated by combining the frequency or probability of threat on an annual basis with the single loss expectancy:

ALE = SLE x ARO

The main disadvantage of the quantitative analysis is that despite its accuracy, it is difficult to be done properly. Usually there is no general agreement among equipment manufacturers and risk analysts as regards the types of threats and the frequency of their occurrence. Another serious drawback is that while the calculations of the quantitative analysis are quite simple and clear, the process of data collection is quite complex and insufficiently clarified in terms of methodology. That is why many corporations and security experts rely on another method of risk assessment, namely the qualitative analysis.

Qualitative Risk Analysis

Qualitative risk analysis uses the experience and intuition of those who know best the company's assets. This means that knowledgeable people will use their experience to assess what risks threaten the assets and what measures should be taken against them.
The process of gathering information by experts is called **Delphi technique**[9].

The advantage of this method is that it is flexible and easy to implement. Trusting the experience of those who work daily with assets, an appropriate regime of countermeasures can be planned.

Since qualitative analysis depends largely on the judgment of people, it is more effective when the number of specialists involved in this process is greater. This method can use questionnaires with grades from 1 to 10, interviews, group meetings, detailed hypothetical questions, and in general, any technique that allows people to express their views in terms of threats and risk reduction. The process of qualitative analysis is similar to the process of quantitative analysis, but it does not use numbers and formulas. First the assets to be protected are identified. Then there is brainstorming to identify the potential risks that threaten each asset. Finally, the group decides what the most effective reaction to each risk is.

In qualitative analysis calculations are simple, but it, in turn has several disadvantages. The main problem with it is the subjective nature of the analysis and the lack of objective information about the value of assets.

Combining quantitative and qualitative risk analyses

Depending on the organizational structure, a corporation's management may by satisfied with the results of the qualitative analysis. Some organizations, however, have to justify their costs, so they need the information that the quantitative analysis produces. In practice, elements of both analyses are used.

In determining the exposure factor (EF) and the single loss expectancy (SLE) there is always a qualitative element present in the quantitative analysis. Few institutions can authoritatively predict that EF is not 0.65 but 0.50 or that the annualized rate of occurrence (ARO) is 0.3, not 0.4. These numbers are defined more precisely in the process of discussion with experts whose opinions are determined by their own intuition and experience.

Conclusion

The security risk analysis is fundamental to the security of any organization. It ensures that the control and the expenditures made are fully commensurate with the risks faced by the organization. In this way an optimal choice is made as to which risks are to be addressed and respectively the steps for eliminating them.

References

1. Threat, vulnerability, risk, http://www.threatanalysis.com/blog/?p=43
2. Stoilov, E. Upravlenie na mrezhovata sigurnost. Sistemi za otkrivane na narushiteli. Sofia, 2011.
3. Avtomatichni informatsionni sistemi – riskove i zaplahi kam softuera. http://www.hightechbg.com/avtomatichni-informacionni-sistemi-zaplahi-software
4. Sigurnost na avtomatiziranite informatsionni sistemi i mrezhi. Akreditatsiya na sigurnostta na voenni KIS. CIO, 2004, issue 5.
5. Landoll, D. The Security Risk Assessment Handbook: A Complete Guide for Performing Security Risk Assessments. Texas, USA. CRC Press. 2011.
6. Vose, D. Risk Analisys, A Quantitive Guide, Third Edition. England. 2008
7. Stewart, J. M., M.Chapple, D.Gibson, CISSP: Certified Information Systems Security Professional Study Guide.
8. Gibson, D. Managing Risk in Information Systems. USA. 2011.
9. Kim, D., M. G. Solomon. Fundamentals Of Information Systems Security. USA. Jones and Bartlett Learning. 2012.

Decision Rules in Agent Based Modeling: Spatial and Economic Dependencies of Crime

Georgi Penchev

UNWE, 1700 Sofia, Bulgaria
Department "National and Regional Security"
gpenchev@e-dnrs.org

Absract. The paper reviews some of more prolific current thoughts in research on criminal activities. It considers the significance of spatial dimension and dependencies of crime in general. The use of Agent Based Models is proposed as possible tool for analyses of such dependencies. The minimal set of rules is suggested. Some graphical representation of crime activities is given as mean for verification. The paper is focused more on presenting a new arguments and thoughts about crime models than to comprehensive presentation neither of literature on crime, nor to Agent Based Models.

Keywords. Economics of Crime, Spatial Analysis and Statistics, Agent Based Modeling.

1. Introduction

The crime is probably one of the most studied human activities. It has many aspects and appearances. The classification of crimes is enormous job by itself. The current report is not aimed in comprehensive study of crime and its roots. It is not aimed even in description and classification the existence literature. Even more, the report does not present the Agent Based Models of crime.

The report has a very modest task to discuss in broad terms the prolific present ideas for existence of relation between economic conditions described as GDP, unemployment, income, social and gender structure of population, etc. and crime rates.

The consideration is made by adding some arguments about current modeling of crime, based on variables which can be treated both as cause and consequence of criminal activities. The spatial dimensions of crime are also discussed. Based on these considerations a minimal set of rules for Agent Base Model is defined and initially tested.

The paper can be considered as a first attempt to define new more practical view to crime study and definition.

2. Crime, Economy, Social Conditions, Counter Action

The crime is a phenomenon which can be considered from many aspects. Probably the first aspect is the legal aspect, but if we look closer to historical development of legal system and legal definition of crime we can see the differences in definitions and systems during the different time periods. Even more, if we look at different countries we can see distinctive differences in measurement, definition, treatment, punishment, etc. Thus we have to look closer to the social, cultural and other aspects of historical and social development of the crime. Perhaps considering these significant social and cultural factors we can find the reasons of crime. But there is crime both in most developed countries, as well as in least

developed countries. The crime present as significant social and legal phenomena in countries with significant social, cultural, religious and legal differences.

Therefore we can suggest that there are more fundamental reasons and factors for crime. It is reasonable then to look at idea of existence of more profound choice with economic roots. Such kind of individual choice can be based on substitution between legal and criminal activities. The reason is simple. If the profit and benefits of criminal activities exceeds the profit and benefits from legal activities, then the individual probably will choose crime activities. [1] The cultural, religious and social properties of the individual's environment will cause the different elasticity of the substitution. And these parameters are easily measured with variables such as crime rate, crime per individual or more complicated variables as prison population on one hand and GDP, GDP per capita, unemployment, social segregation, poverty and income distribution on the other.

Undoubtedly, the suggestion for legal/crime choice has huge explanatory power. It can be used for explanation of the connection between economy conditions and crime rate. If the economy is in recession or depression and the legal opportunities are reduced by crisis then the crime become as more seductive as the economy condition are worsened or the level of poverty and segregation of some social group are risen.

Unfortunately, this strong reason is not enough to create a model of the crime. There are a lot other dependencies which have their influences on crime. The crime counter measures are one of them. There are attempts to implement the counter measures within the crime/legal activities choice as an external variable. The counter measures can be treated as factor which makes the cost of crime activities higher, thus balancing the choice in new level. But, it is reasonable to claim that the crime rate and counter measures are related and it is hard to define which of variables – crime or countermeasures is independent. The increase in crime will lead to increase in protection demand. On the other hand, the level of police presence and activities will influence significantly the crime rates.

The quality of counter measures is other related variable. The high quality counter measures in time of high crime rates can be deteriorated by the number of cases and the time needed for investigation, prosecution, etc. Thus the excellent working court system can be spoiled with crime rate and prison population increase.

The theory of economic choice of crime can be considered as very conditional even from economic point of view. During the time of crisis the demand for stolen goods also can be reduced, thus lowering the attractiveness of crime. We can argue that if all income groups in society are equal affected then the "market" of stolen goods will be shrunken. Otherwise – the decrease in markets of stolen goods can be compensated if middle and higher income groups are less affected from the crisis. Also, considering economic roots of crime we can look at cost of crime activities and crime infrastructure as money laundering and other activities needed for cleaning of stolen goods or money laundering. [2]

The consideration given above about difficulties of creating one unified economic model of crime is confirmed by many researches.

3. Crime and Crime Studies

The crime as phenomena is studied in many researches from many points of view and through many approaches. The variety of crime studies is so vast, that there are many articles dedicated to summarization, comparison and classification of economic and social crime studies. [3]

There are also many studies ordered by different governmental bodies – national and local. Usually these studies are dedicated to estimation of crime cost from society point of

view. Many of other studies considered the effectiveness of counter measures and prevention polices such as cost and benefits of additional police officers employment, regulation changes and other. [4]

It is worth to look at studies of social crime costs. The world wide review of literature prepared for office of Attorney General & Justice, New South Wales, Australia shows that there are serious variations of crime costs estimation among authors in regard of same type of crimes. As example the cost to society of homicide is estimated from 2 million dollars to about 19 million dollars. [5] Nevertheless there are certain similarities in authors' opinion in regard to comparison among types of crime – the crime against the person are 10 to 12 times more expensive to society than crime against the property. Interestingly, current report of World Bank to Jamaica, suggests to the Jamaican Government to take emergency measures against homicides in order to support GDP growth by avoiding substantial social costs of these crimes. Costa Rica's economic growth received by considerable homicide rate decrease was given as an example by the report. If we consider this suggestion in above mentioned difficulties of identification of independent variables in crime models we can add one additional relation, e.g. one additional difficulty. [6]

Many international and national authorities and research centers are caring out crime statistics and studies. The topics are related to comparative studies among countries, impact of migrants to crime rates, international and cross-border crime.

In regard of the subject of this paper – we can argue that the classical point of view prevails – the higher level of unemployment, the worsened social and economic conditions mean inevitably more crime.

4. Crime, Data and Aggregation

The research on current economic crisis, funded by USA Congressional Research Service, which main objective is to find necessarily balance between police workforce and violent and property crime shows some contradictory results in relation to the classical economic view to relation growth and conditions – crime. According to the report many researchers studied only these crises where the relation is positive. The more broaden view – from 60-es to 2009 shows times with positive and times with negative relation between crime and economic condition (growth or downsize in studies' terms).

In addition to these observations the report suggests that the aggregation of data on national levels blurs the situation on state and counties levels and suggests not trying to find a clear theoretical model of crime, but to find the reasons whether or not to employ more officers. [7]

Many EU countries, especially the developed member states have excellent statistics on crime. Germany, France and UK have special statistical units responsible for crime monitoring. Nevertheless their reports are aggregated and published in formats that hardly can be used for studies. National statistics of mentioned countries give more or less detailed excerpts of this data. There are no data which is spatially (geographically) referenced.

Before 2014 EUROSTAT provided very few variables of crime statistics. Since the beginning of 2014 the crime data is given in time series for period 2003-2013 based on national level. There are georeferenced data, on EU's NUTS level 3.

Even in this format the data shows that there are many differences between EU's member states. The relation GDP-crime rates vary significantly as Table 1 shows. There are member states showing positive or negative strong, middle or weak relations.

AT	-0.925622
BE	0.856185
BG	-0.621983
CY	0.353865
CZ	-0.701521
DE	-0.894958
DK	-0.211402
EE	-0.749612
EL	-0.007676
ES	0.947117
FI	-0.019863
FR	-0.694434
HR	-0.835044
HU	0.317575
IE	0.130307
IT	0.841865
LT	-0.735798
LU	0.688199
LV	-0.239313
MT	-0.739291
NL	-0.781892
PL	-0.950318
PT	0.246441
RO	0.60833
SE	0.593707
SI	0.489626
SK	-0.869334

Table 1: Correlation coefficient between GDP per capita
and Crime Recorded by Police for the period 2003-2012
Source: EUROSTAT Databases on National Accounts and Crime and Criminal Justice
(calculated by the author)

5. Crime, Models and Space

Considering the above mentioned arguments leads as to questions such as: Is it possible to create stable, detailed, verifiable model of the crime and its economic roots? Is this search is only academic endeavor with no practical dimensions? Are these propositions true only when the rule "ceteris paribus" is engaged and have only partial explanatory power?

Probably, most of answers are positive. It is hard to believe that crime as phenomena can be explained as stable relation with so many stochastic and unstable variables as cost and benefit of crime, number of police workforce in some moments, etc. The process of choice whether to do or not to do crime is discrete and is dependent of so many variables which in different conditions may or may not be sufficient. [8] In addition it is hard to distinguish dependencies among variables – which is cause and which is consequence. Many models also suggest too reasonable behavior for criminals. The last, but not least too few of crime

studies consider the infrastructure of the crime – the demand side of the economic model – people or organizations which provide services related to stolen goods resells and money laundering. Without good and reliable infrastructure crime will hardly pay.

We can add to crime infrastructure (considered in above mentioned terms) and possibilities for crime 'workforce' recruitment. This means that groups of experienced criminals have to have close contacts with possible recruits.

The infrastructure of the crime as every infrastructure has its spatial dimension and we can argue that this dimension is significant for the crime. The spatial element of crime is crucial for better explanation of the relation between economic conditions and crime. If there are segregated by income territories the lower income parts can be used as good base for crime recruitment. If there are many places which are hard to be monitored frequently enough then the criminal services can flourish on this places. The presence of infrastructure can explain a market-like behavior in criminal activities.

Therefore we can suggest that the crime model has to implement spatial elements and variables on first place. But if there are spatial elements how they can be aggregated? Is it possible or preferable? It is possible to aggregate both the crime cases on given region and conditions in the same region. But it is not preferable. By aggregating cases we can lose important information about spatial disposition of elements and their spatial dependencies.

Therefore the most explanatory models could be considered the most detailed spatial enabled models, as it is suggested in before citied study of Congressional Research Service[5].

If we look at practical use of aggregated models we hardly could find a good practical reason to use such kind of models. They show only the increase or decrease of crime rate or other related variables. What can as an example central government does with such kind of models? To hire more policemen, judges and prosecutors? But where to dispose them? In which region? With what objectives?

The answers on these questions cannot be expected to be derived from highly aggregated models. The local models with well explained spatial dependencies could be much more useful and practical.

Summarizing, the crime should be modeled as economic and spatial phenomena on the lowest possible level with no aggregation. As lowest possible level can be considered the urban or metro areas in which market-like criminal activities can be organized and sustained. The models should be aimed mainly toward identification of alternatives in decision making related to crime counter-measures. The influences of national and international crime to the studied urban area can be modeled within terms of local spatial dependencies.

6. Agent Based Models and Agent Rules

One additional rule should be employed to define the better model. The model should not be defined as one equation, but as many independent conditions (equations) which have to be used in discrete decision making process which depends of current conditions. Such kind of model can avoid the vague interdependencies between variables. Such kind of models can be created by using of Agent Based Modeling (ABM). [9]

Here we will not consider all characteristics and use of ABM. The aim of paper is only to suggest type of agents and rules for their behavior. To create a very simple understanding for current analysis needs we only define the agents as computer programs operating under one or few simple rule. This computer agent is put in environment with constant or changing conditions and it decisions are autonomous. There can be many different species of agents. The result of interaction of agents with one another and with

environment is what is observed and analyzed. The ABM models are widely used in current economics in so called Complex Systems Studies. The idea behind these models is to reproduce the complex behavior on macro level by modeling the simple rules and behavior on micro-level. [10]

The ABMs have their limits. One limit is the computational power which limits the number of agents. The second is the software – not each ABMs software supports use of georeferenced data or spatial analyses.

The AB model can be very detailed – we can create many species with hundreds of thousands of agent each. We can give a lot of instruction of these agents. The species can reproduce in details the social structure of a city or to present different ranks in police or in criminal world.

The level of details can be defined according to objectives of the modeling, but usually more details mean more costs and time. We will try to define minimum number of species and agents and to consider some rules for agents as follows:

- Level of income in residence area, unemployment, social structure, e.g. parameters related to economic condition and quality of life in the neighborhoods. This parameter is variable and should be implement as specie of agents;
- The cost of counter measures – special variable relating to the distance from police department or other anti-crime agency and the number of force to some point in available space (in ABMs usually it is called 'scene'). The cost can be presented as parameter of streets. This is also variable – the places can be changed. And roads should be implemented as specie.
- The population – the number of inhabitants with different levels of income or other related properties. The properties can be changed according to simulated conditions – growth or downsize.
- Arrests, prisons and courts – species which represents capacity and speed of investigation and prosecution.
- Stolen goods resellers, time for cleaning goods and for resell.
- Crime demand – variable, based on possible market share of stolen goods and economic conditions.
- Propensity to crime – variable calculated for each agent of population which represents the risk to engage in criminal activities, based on income level, closeness to low level neighborhoods and crime demand.
- Private demand for security – variable of agents from population, which describes the demand for private security measures as alarms, private police, etc. It depends of agent's income level. Higher the income level means higher demand.
- Cost and benefit of crime – benefits can be calculated on the base of possible victim – higher income of housing area and the level of crime demand. The cost can be represented by the cost of counter measures, private demand for security and operational costs.
- Operational cost of crime can be represented by some standard costs for planning and execution of crime. It can vary by level of private security and counter measures costs (more measures to overcome – more costs), etc.
- Change of economic conditions – can be implemented as species which are changed according rules based on time and probability.

The most of proposed variables can be predefined by statistical calculation or can be based on public opinion studies.

7. Some Spatial Examples

In order to achieve some verification of theoretical thoughts presented in this paper the following phenomena should be observed in urban areas:

- Spatial separation on low and high income neighborhoods of housing area;
- Closeness of police department to crime activities;
- Distribution of crime activities in zones with higher distance from police departments;
- Lower level of crime activities in higher income areas in regard to higher level of private security.

The examples are based on data gathered from web-sites of three US cities – Denver, New Orleans and San Francisco – three urban areas with 300 to 500 thousands of population. The georeferenced data used by web sites is as follows – crime as it is recorded during year by 911 calls and plan of streets. The income data on tract level (tract is relatively small administrative unit of county, the states are consistent of counties) was received by US Census 2010 data through American FactFinder and georeferenced by maps provided by TIGER System.

The 911 calls are filtered by category of 'burglary' and by time period of one month. There are many other types of recorded crimes, but probably the burglary is most dependent from above stated rules.

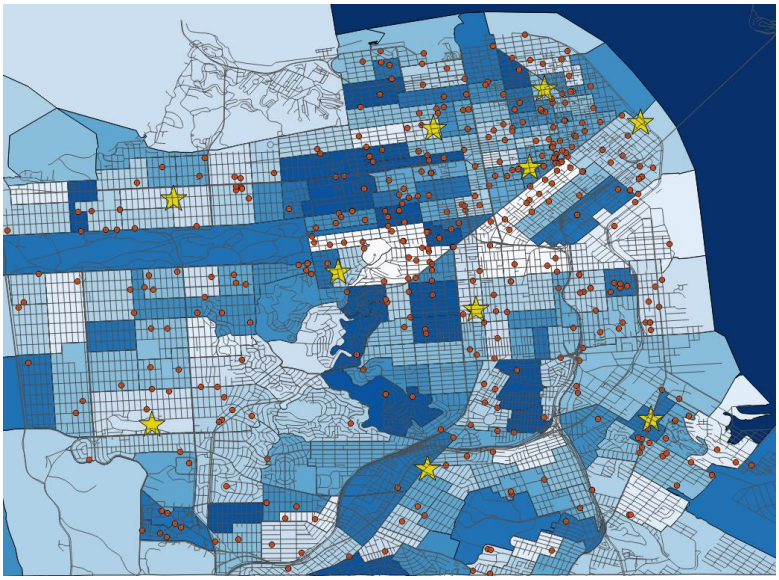


Figure 1: Crime of category "Burglary", San Francisco January 2012
Sources: San Francisco Data (https://data.sfgov.org/) and US Census 2010

On Figure 1 is shown the burglary distribution for January 2012 in San Francisco and places of Police Departments (PDs). The crimes are distributed mostly in north-west part, together with 3 PDs (yellow stars). The tracts are with middle to high income. In highest (dark blue tracts) and lowest level income tracts (white and pale blue colored tracts) crimes are much rare.

The same picture can be observed in other two cities.

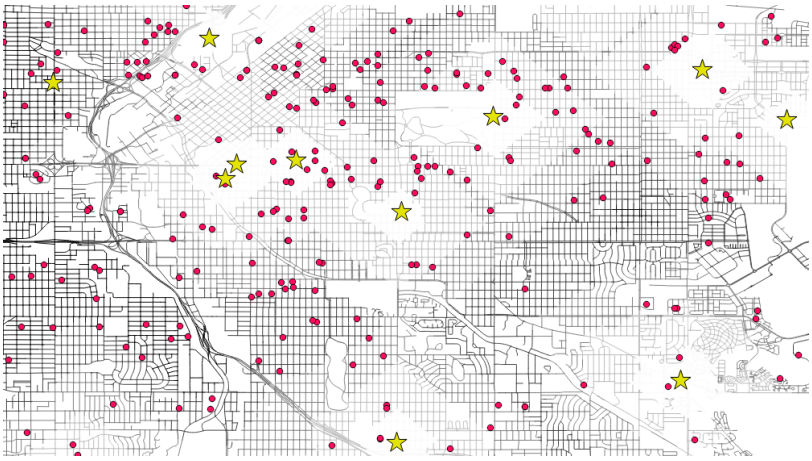


Figure 2: of category "Burglary", Denver January 2008
Sources: Denver Open Data Catalog (http://data.denvergov.org/) and US Census 2010

Figure 2 shows the distance to PDs in Denver – the grade of street colors shows closeness – more pale color means closer distance. Most of burglaries are in more distanced from PDs places.

The same situation is observed in New Orleans and San Francisco.

The observations are pure graphical, but distinctive and deserves future detailed statistical analyses.

8. Conclusion

In conclusion it can be summarized that literature and studies dedicated to crime and crime modeling is a very broad area, which envelopes many approaches from different branches of science and especially from economics. The search of relationship between economic conditions and crime in general is very powerful and has its explanatory power.

Nevertheless, these ideas suffer from some deficiencies which are caused mainly by aggregation of data and omission of very significant spatial dependencies of crime.

The models of crime can find more practical application if they are concentrated on decision making than search of full explanation of all aspects and dependencies of criminal activities. The use of Agent Based Modeling is one possible way in such direction.

The paper have to be considered as initial work in more broad subject on crime and its spatial analyses, but even such it gives some promising results.

References

1. **Freeman, Richard B.** The Economics of Crime. [ed.] Edited by O. AshenJ~lter and D. Card. *Handbook of Labor Economics.* s.l. : Elsevier Science B.V., 1999, Vol. Volume 3, Chapter 52.
2. **McCrary, Justin.** Dynamic Perspectives on Crime. [ed.] Bruce L. Benson and Paul R. Zimmerman. *Handbook of the Economics of Crime.* s.l. : Edwar Elgar Publishing Limited, 2010.
3. **Steven D. Levitt and Thomas J. Miles,** Economic Contributions to the Understanding of Crime. *Annual Review of Law and Social Sciences.* 2006, Vol. Volume 2, pp. pp. 147–164.
4. **Bruce L. Benson,** The Allocation of Police. [ed.] Bruce L. Benson and Paul R. Zimmerman. *Handbook on the Economics of Crime.* 2010.

5. **Webber, Andrew.** Literature Review: Cost of Crime. [Online], http://www.crimeprevention.nsw.gov.au/agdbasev7wr/_assets/cpd/m660001l2/cost%20of%20crime%20literature%20review.pdf
6. **Observer, Jamaica.** Cut crime and GDP could increase 5%, says World Bank. [Online] April 14, 2011. http://www.jamaicaobserver.com/news/Cut-crime-and-GDP-could-increase-5-percent-says-World-Bank_8667965.
7. **Finklea, Kristin M.** *Economic Downturns and Crime.* s.l. : Congressional Research Service, December 19, 2011.
8. **John Roman and Graham Farrell,** *Cost-Benefit Analysis for Crime Prevention: Opportunity Costs, Routine Savings and Crime Externalities.*, Volume 14 , Crime Prevention Studies,, pp. pp.53-92.
9. **Charles M. Macal and Michael J. North,** Introduction to Agent-based Modeling and Simulation. [Online] November 29, 2006. http://www.mcs.anl.gov/~leyffer/listn/slides-06/MacalNorth.pdf.
10. **Robert Axelrod and Leigh Tesfatsion,** Agent-Based Modeling in the Social Sciences. [Online] http://www2.econ.iastate.edu/tesfatsi/abmread.htm.

Data Sources

1. US Census Bureau, American FactFinder – Census 2010 Data for Income on Tracts of Orleans Parish, San Francisco and Denver County - http://factfinder2.census.gov/faces/nav/jsf/pages/guided_search.xhtml
2. US Census Bureau, TIGER – Geography, TIGER/Line® Shapefiles Pre-joined with Demographic Data, Shape files of Tracts of Orleans Parish, San Francisco and Denver County, https://www.census.gov/geo/maps-data/data/tiger-data.html
3. Denver Open Data Catalog, Shape files for street centerlines, Crime and Police Departments, http://data.denvergov.org/
4. San Francisco Data, Shape files for street centerlines, Crime and Police Departments, https://data.sfgov.org/
5. Enabling New Orleans, Shape files for street centerlines, and Police Departments, Georeferenced file in csv format for Crimes, https://data.nola.gov/
6. EUROSTAT, GDP and main components - Current prices (nama_gdp_c), http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_gdp_c&lang=en
7. EUROSTAT, Crimes recorded by the police (crim_gen), http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=crim_gen&lang=en

Application of Heuristic Methods for the Assessment and Classification of Threats to Information Security

Nedko Tagarev

Abstract : This article, is related to the use of heuristic analysis methods that are undervalued in the analysis and assessment of threats and risks to information security . Many experts in the field of information technology , even denying the analysis and planning , and aspects of information security . This contributes to the direct work of IT professionals with customer in relation to the crisis recovery systems . Overall, it can be noted the absence of valuation of assets and values, that depend on the security of information. From the other point of view, often, the difference between threat and risk still is misunderstood. Actually wrong understanding of the origins and consequences of threats to the assets and values leads to the suspension of the operation of information systems. This publication shows the application of expert analysis methods in terms of information security.

Keywords: security , information, asset value , analysis, expertise and methods

Analysis by scenarios of system security management of information

Overall , scenario planning is a technique by which stimulate the presentation of different perspectives or visions of the future in a particular area to be able to make better predictions of the development scenarios through her analysis is the process of analyzing possible future adverse events regarding ISMS . It enables the construction of alternative versions of existing systems for information security . This scenario planning provides the ability to monitor outcomes and possible future states of information assets. Unlike projections methods of analysis, the analysis with scenarios using extrapolation of the past , in case when it is not expected events of the past to influence the future development . By analyzing several scenarios show possible future changes in ISMS*(Information security management system)* . This is useful for generating a combination of optimistic, pessimistic and most likely scenario . [1] Experience shows that is necessary to use at least three scenarios for further discussion and decision-making . Analysis through scenario is intended to improve the process of decision making regarding information security . Moreover , it allows to take into account the results and consequences of decisions . This type of analysis in IC, allows to analysing particularly the relationship between specific interrelated attacking agents and protective mechanisms in ISMS.

With analysis through scenarios in the field of information security, we have the opportunity to change the usual, using of statistical laws. This occurs due to the possibility that decisions are taken outside the limits of the standard statistical setting. The bottom line is not allowed chance. They are based on arbitrary assumptions, and there is no possibility of setting expected values.

Process analysis regarding the amendment of the assets in the management of information security

The process is a sequence of operations transforming the input resources into a desired end product . "[1] In a complex process consisting of a number of actions or those

performed in a long period of time , there may be several distinct intermediate results of modified resource input or change the system from which flows studied process. [1]

Practice has shown that the application of the process approach for the analysis of the security of the information leads to a number of positive effects , including:

• Reduce costs .

• Increase the effectiveness of protective mechanisms.

• Increasing the level of manageability , including improve the ISMS , transparency in the management of information security , speeding up procedures for decision-making concerning the improvement of the defense mechanisms.

• Reduce the impact of the human factor in the threats and risks to information security .

Process analysis is primarily used to analyze the change in asset values and information security . Any change consists of processes. They are designed to change the value by transforming the resources assigned to the inputs of the system of final output. The output of the system stands the future state of the changed or modified asset. Processes can have a significant impact on the security of the information. Improving a defense mechanism can improve the functioning of the organization as a whole.

The first step in improving the analysis process . It is necessary to understand the operations , the relationship and the values of the corresponding parameters. Analysis of the process includes the following tasks:

1) Defining the limits of the process. Mark the input and output, and input and output values of the assets and values of information security ;

2) Construction of a flow diagram . Activities, processes and their connections ;

3) Determination of the capacity of each step in the process. Calculation of the quantitative indicators of the change in the asset or value ;

4) Identification of the busiest process as a separate step with the lowest capacity;

5) Assessment of additional restrictions. Determine the quantitative impacts;

6) Use the analysis for operational decision-making and the establishment of the need to change the security mechanism.

Process analysis is represented graphically by a flow diagram . The borders of the process is determined by the entry and exit points that define the inputs and outputs of the process. Flow chart is a valuable tool for understanding the process .

An example is shown in flow chart (Figure.1.) Three successive operations:

Figure.1. Flow diagram of the process analysis [2]

Symbols in the flow diagram are defined as follows:

• Rectangles represent tasks;

• The arrows represent flows. Streams in this study include the flow of the attack and its power. The flow of attack involves the knowledge and resources of the attacker's agent , which determines the strength of the attack.

• Converted triangles represent repositories . These are the cells that are used to represent the input and output values and status of assets information security .

In the flow diagram tasks in series one after the other , are consistent and tasks that are parallel are performed simultaneously.

When building a flow chart should be careful of " traps " that could lead to establishment of such a flow diagram that does not represent a real change in the assets. In deciding managers are guided by indicators - price, quality, flexibility and speed.

In the process of analyzing the information security these indicators include the following:

- Capacity of the process.
- Capacity utilization.
- Percentage of flow.
- Current time.
- Cycle time .
- Time attack.
- Standby time.
- Work in process.
- Default.
- Specific content of the work.
- Specific use of human resources.

The main purpose of the process is to improve the analysis process . Usually improvements are related to cost, quality and time. The analysis of costs and benefits carried out to determine whether the process of change is worth the investment . After all, very often NPV(Net present value) determines whether the process " improvements " really lead to that.

Expert analysis methods

Expert analysis methods are part of the class of intuitive estimation methods. They are always based on the opinion of one or more experts. Because of this feature , the reliability of the estimates depends critically on the competence of the experts whose selection is becoming one of the most important elements in the organization of such estimates. [3] The choice of experts, predict a spot on the immediate guess work , so have to be carried out with caution. [3]

Brainstorming [4]

Brainstorming (BS) , is a group technique for generating new and useful ideas in making decisions and increase creativity . This technique can be used as a tool to define the task on which they work , identifying specific problems , finding solutions and sample testing whether they are reliable . [5]

Three positions of the participants in this method:

- team leader ;
- a person who recorded and
- team members .

The team leader must be a good listener. In the beginning, he helps the participants. He must give to them the reasons why it is necessary to proceed to brainstorming, and to increase their participation. Also, this leading person should set the ground rules and to manage the process of generating ideas .

Writer participant record all ideas. This role can be performed by team leader.

The participants in the team must be not less than 5 and not more than 10 . It is best to have six or seven .

To objectively study , there are rules for the meeting , which must be strictly observed.

There are three important factors that determine the success of a brainstorming session .

1) First, the group must strive to create a large amount of ideas , as this increases the possibility to find the best possible solution.

2) Second , the participants has certainly not help to criticize ideas as they share, as negative thinking of one member of the group can make others less willing to participate. Thus lead to failure the whole process.

3) Third, the group leader should create a positive environment for BS and guide, the creative energy of the other members in the same direction .

6-3-5 method

6-3-5 method encourages a process of problem solving with generation of new and unusual ideas in a group of people. In Method 6-3-5 participants receive six equally large sheets of paper. They were divided into 3 columns and 6 rows with 18 cans. Each participant completed on the first row three ideas. The sheets are exchanged after a certain period of time - depending on the severity of the problem –from 3 to 5 minutes , in a clockwise direction . The next attempt is to answer the above ideas , and complements them and develop.

The name of the method is the result of optimal six members of the group that produced three initial ideas and then five times, changing leaves and the first three or the next three derivatives ideas for development (6 participants every three ideas , 5 times) . Often the " 5 " in the title of the method is associated with the processing time - not more than 5 minutes.

This method within 30 minutes, leading to more than 108 points: 6 participants x 3 x 6 rows ideas.

Application of heuristic methods in analysis of IC

The key factor for the successful functioning of the methods is the time that is used to analyze the information security . The author's opinion is that usage of this methods is sufficiently fast and easy for the staff responsible for information security . The main obstacle comes from the speed with which define vulnerabilities and threats to information security .

In the thesis proposes the author proposes two approaches for solving this problem . The first is related to the proper functioning of the early warning of emerging vulnerabilities and threats to information systems. The commencement of the proper functioning of the ISMS the author decided the issue at a second approach using expert judgment .

For the purpose of the thesis of the article for analyzing information security expert survey was conducted by the methods of brainstorming and 6-3-5 method . Brainstorming is used to generate ideas about the major threats and vulnerabilities. Based on this study were generated following threats focused on the information security:

- Threat of natural disasters - earthquakes , floods, too high or low temperature ,
- Social engineering
- Unmotivated staff;
- Accident - flooding , power failures , leaking chemicals
- Espionage
- Lack of risk analysis ;
- Act of terrorism - bombings ,

- Failure in the system
- Lack of standardization;
- Threats from malware;
- Faults in the creation of the system
- Lack of legislation.
- Intentional hacking
- Untrained staff ;
- Lack of control over safeguards ;
- Staff who intentionally fails to comply with the security rules;

This treats is used in analysis by scenarios.

This expert method for analyzing and forecasting is a good starting point for the analysis of information security. To confirm the results of the brainstorming author uses 6-3-5 method, which is suitable for prioritization of identified threats. This is necessary for organizations with limited financial resources. Based on 864 reviews generated 80 different ideas.

Number of unique idea threat / risk many times as an idea Proportion of all ideas Share in percent

1 Terrorist attack / sabotage 68 0.08 7.87%
2 natural disaster (flood , earthquake, tornado (1)) 58 0.07 6.71%
Hacker attack 3 43 0.05 4.98%
4 Industrial Espionage / Social Engineering 42 0.05 4.86%
Malware 5 39 0.05 4.51%
6 Negligence and / or incompetence of staff 37 0.04 4.28%
7 Emergency (fire, electric shock , flood, flooded) 37 0.04 4.28%
8 Failure IS 31 0.04 3.59%
9 Weak / insufficient protection 27 0.03 3.13%
10 spying devices near the critical information / interception 24 0.03 2.78%
Unfair competition 11 23 0.03 2.66%
12 Lack of motivation in employees 23 0.03 2.66%
13 Poor recruitment / hiring now 20 0.02 2.31%
14. Leaks 14 18 0.02 2.08%
15 Missing / outdated regulations 18 0.02 2.08%
16 Introduction of a new information system / new technology 18 0.02 2.0

Factor to influence the risk

For the assessment of the impact (impact factor) will be used qualitative criteria , which is shown in ISO 31000 . Influencing factors will be measured by whether there is a change in the status of the asset. If there is a change , the value is 1 if there is no change , the value is of 0. According to the impact assessment through qualitative criteria has the following categories:

• Negligible - There would be no impact if the threat occurs ;

• Small - There is a little effect. Requires no special effort to repair ;

• Significant - has complicated effects as a result. Suggests paying attention to the side of a few employees . The result may be affected limited amount of data . Requires time and resources for recovery;

• Serious - Violation of the interests of the partners and employees. As a result, much of the confidential information may be made public ;

• Severe - Disclosure of sensitive organization data. Continuous violation of the organization's reputation . Can cause company closing;

• Demolish - damage may have violated the reputation and / or lead to loss or disclosure of private information . Efforts are also great resources for recovery;

These criteria represent the relative weight of the violation of any one of the values. The relative weight given by expert judgment .

For the assessment of the expectation will be used criteria, which is shown in ISO 31000. According to the estimation of the probability with qualitative criteria with the following categories:

• Negligible - usually happens -0, 05
• Very low - Sometimes two / three times every five years -0, 1
• Low - Once a year or less 0 - 0.1-0 , 25
• Average - Once in six months or less 0 - 0.26-0 , 50
• High - Every month or less - 0 51 -0.79
• Very high - many times a month - 0 0.8-0 , 89
• Ultra - Multiple -day -0.9 -1

Process analysis allows drawing a more accurate assessment of the effects through qualitative criteria. Accordingly, each of the three assets being tested values, information security, economic values and material values, there are three indicators that change.

• Values for information security - confidentiality, integrity, availability,
• Economic values - finances, reputation, legal consequences;
• Material values - tangible assets, human sacrifices security.

In negative change any of these parameters is set to 1,. Accordingly, the scale is measured from 0-9. The values and assets should give relative weights. This is achieved by expert judgment.

Defined by the author material values are of most importance for the security of information, with 20 of the 50 references, or approximately 40%. Second rank is economic values with 17 responses or 34% accurate. Finally, there are the values of information security with 13 references, or approximately 25%.

For purposes of the survey questionnaire includes determining the importance of the assets of any value. The author does this in order to determine the relative weighting of each asset. It is given on the basis of an expert study conducted with 50 experts in the field of information security.

Most importance has the confidentiality with 22 references, or approximately 43%. Second is the consistency of information with 15 replies or just 30%. Third place take accessibility with 13 references, or approximately 27%.

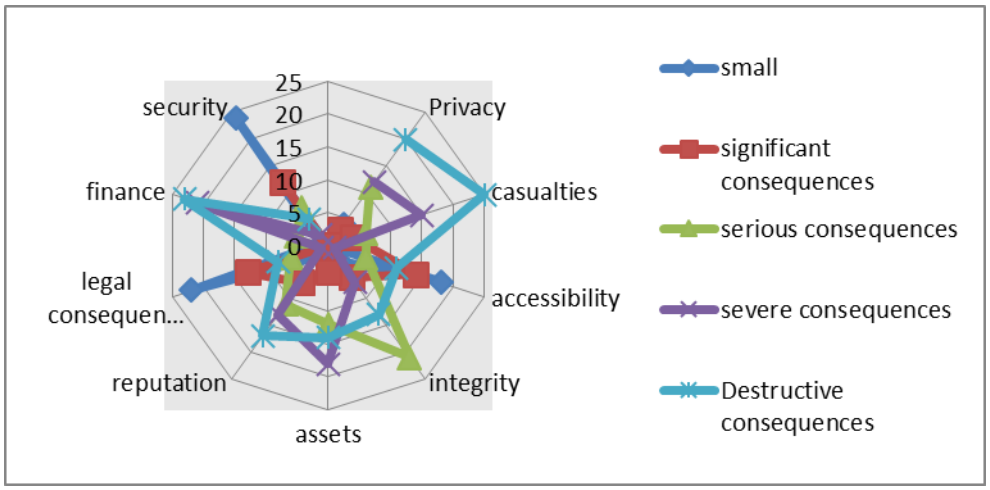
First in importance of the asset is determined finance with 24 references or just 48%. Second is the reputation of 20 references or 40%, and in third place came the legal consequences by 12% or exactly 6 replies.

Results are as follows for the importance of tangible assets on the basis of the study. The first is human life by 52% or 26 references. Second with 22 references and 38% assets, and third with 13 references, 10% is the sense of the security situation.

To complete the survey the author made consensual assessment of the impact of the consequences of violation of information security at the sites of critical infrastructure. The assessment is based on the opinions of 50 experts in the field of information security, interviewed. The survey results are shown in Table 1.

	small consequences	significant consequences	serious consequences	severe consequences	Destructive consequences
Privacy	4	3	11	12	20
Casualties	0	4	6	15	25
Accessibility	18	14	6	1	11
integrity	3	6	21	7	13
Assets	2	4	12	18	14
Reputation	2	7	11	13	17
legal consequences	22	13	6	1	8
Finance	1	0	5	21	23
Security	24	12	7	2	5

Table 1. Expert assessment of the impact in case of assets.



References

1. Aaker, David A. (2001). *Strategic Market Management*. New York: John Wiley & Sons, pg. 108
2. http://www.netmba.com/operations/process/analysis/(07.07.2012)
3. Иванов., Т. „Икономически анализ в публичния сектор", София, 2009 г., pg. 219
4. Sutton, R, BusinessWeek, 26 July 2006
5. Osborn, A.F. (1963) *Applied imagination: Principles and procedures of creative problem solving* (Third Revised Edition). New York, NY: Charles Scribner's Sons
6. http://bglog.net/Obrazovanie/13374 (05.08.2012)
7. Драганов, Х. „Управление на риска", Тракия – М София, 2003г., pg. 16.

A Dual-Layer Data Model for a Scalable Educational Social Network at a University

Venko Andonov[1]

[1] UNWE, 1700 Sofia, Bulgaria
vgandonov@gmail.com

Abstract. The growing significance of the usage of different social networking services in a personal and enterprise context provoked the interest of many researchers to experiment with Educational Social Networks. The standard approach to building an Educational Social Network is using a generic open source solution (like Elgg or BuddyPress) and customizing the user interface and the login system to fit the specific needs. Following such a deployment at the University of National and World Economy (which has a large potential user base – around 20,000 students and more than 500 lecturers), experiments have shown that the classic normalized relational database model of Elgg is not well suited for a significant growth in the user base, the connections and the shared content (the social graph). This paper presents a new dual-layer data model, which attempts to provide a manageable vertical and horizontal scalability. The model is evaluated by comparing the latency of the most critical (for performance, memory consumption and educational usage) operations, both as the social graph grows and with the corresponding values in the original solution. The underlying causes of the differences are analyzed along with the accepted trade-offs.

Keywords. educational social network, data model, scalability.

1. Introduction

The online social networks, also referred to as social networking sites, show continuous growth of the significance and depth in the personal and professional life [1], the scientific community [2] and a wide variety of areas. For example, another emerging branch in the social networking space are the Enterprise social networks – social environments aimed at improving the communication workflows in the enterprise. The integration with the existing software systems is considered a key to a better adoption and improved productivity of the employees [3].

An educational social network (EduSN) is a type of online social network which provides either an end-user software product or a platform and framework for enhancing the learning process for all of the participants in a given formal or informal educational environment [4, 5, 6]. For example, in the higher education, such a system might be an engaging way to promote some desired activities among the students, including, but not limited to: teamwork; professional communication; sharing knowledge and experience; aggregation and evaluation of information sources; presentation of the individual skills; discussions; construction of knowledge. The most significant difference from a traditional e-learning system is the main workflow – it is not focused around the content itself but the connections between the participants or between the participants and the content, formed actively or passively during the interaction with the EduSN. Thus the purpose of an educational social network is not to replace the existing e-learning solutions with a better

alternative, but to give the students and the lecturers an additional environment (which, from their point of view, might be transparent or not), consisting of familiar tools and behavioral patterns.

The standard approach to building an Educational Social Network is using a generic open source solution (like Elgg or BuddyPress) and customizing the user interface and the login system to fit the specific needs. Following such a deployment at the University of National and World Economy (UNWE) (which has a large potential user base – around 20,000 students and more than 500 lecturers), previous work has shown that the classic normalized relational database model of Elgg and other internal information systems is not well suited for a significant growth in the user base, the connections and the shared content (referred to as the social graph, or more specifically: educational social graph). The integration with internal (from the university point of view) systems and databases further increases the latency of a significant portion of the executed operations. The main goal of designing and developing a data model and architecture from scratch is to provide an abstraction allowing for rapid development and deployment with performance and scalability results comparable to the use of graph databases [7, 8].

2. Data Model

A new, dual-layer data model has been designed and implemented with the following goals:

- being a platform which integrates external services and gets integrated into such (both client or server-side) to serve as a behind-the-scenes social tier in the boundless educational experience;
- natural, developer-friendly interface for educational social graph access, traversal and modification;
- abstractions, allowing the extensibility and evolution of the specific data model;
- easily swappable infrastructure components (persistent storage, caching, logging, authentication and authorization, web service stack).

The logical layer is an abstract graph representation of the in-memory, persistent and remote storages and services. The educational social graph is a multimodal, multirelational, directed graph with key-value properties attached to each node and edge. Its physical storage is a subject to multiple adapters, which process a given subgraph and its metadata in order to retrieve or modify the data using the proper set of methods.

The functional layer models the specific educational context and workflow. It includes a set of basic data structures and operations, which can be extended.

There are two main trade-offs included in the model: availability over consistency (in case of a network partition, the availability of the service is considered more valuable unless the specific data model components are set in a different way) and data redundancy (multiple copies of the same data are kept – replicated by keeping the structure or cached with a different than the original structure).

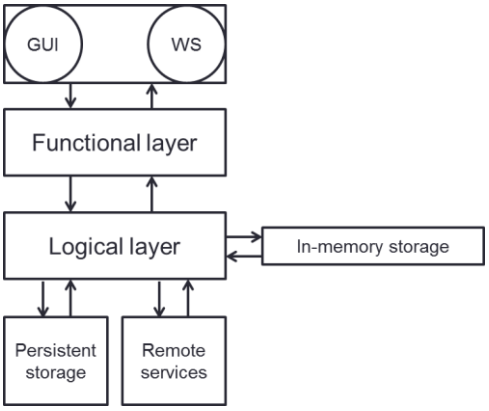


Figure 1: Overview of the suggested framework for design and development of EduSN

a. Data Structures

The data structures in the logical layer are: node (might represent any object or part of an object in the educational social graph – for example: a user, an article, a study group), edge (represents a connection between any two nodes, has a type and a direction), subgraph (a set of nodes and edges between them that together represent a functional object) and collection (a set of nodes or edges, without explicit connections between them – for example, such a collection might be the result of a search query).

In the functional layer, the suggested basic data structures are:

- item – any content, might include text, references to images, video, files, external links; in the context of an EduSN an instance of "item" might be an article, a video lecture, a scientific publication, a discussion thread, etc.;
- participant – any user (student, lecturer, prospective student) or external application that might actively initiate creating, updating or removing items and relationships;
- relationship – it can be any link between two participants, a participant and an item or between two items (in terms of social networking these might be a friendship, liking content, authoring a blog post, rating a book, etc.);
- audience – contains a list of participants and sub-audiences that together form the target set of a shared content item; it is used to enforce the privacy of the shared content.

The functional layer is where the specific data structures can be defined in terms of the available data structures and operations of the logical layer.

b. Operations

The operations in the functional layer are categorized according to several criteria – expected educational usage, the interaction with the logical layer and the provided API (this allows for precise measurement, analysis and management of the performance of the system). The interaction with the educational social network is happening through the exposed API and its relevance for the classification of the operations is based on analysis of the features usage: the frequency of the execution of a given feature; the interaction graph (as constrained by the hypermedia API); behavioral patterns of the end-users; expected patterns of internal and external applications using the educational social network as a service. These indicators can be used for preliminary assessment of the expected usage of the operation and further optimizations.

As precise preliminary data cannot be obtained, to assess the potential usage of the operations, the following sources are considered:

- statistics from the currently operational educational social network, built with Elgg;
- scientific publications related to research on the usage of any type of social networking services, the used features, client devices, etc.;
- existing research on the properties of the social graphs;
- existing graphical user interfaces (web, mobile) of generic, enterprise and educational social networks;
- existing external (third-party) applications for social networking, both in personal and enterprise context.

3. Implementation

The implementation should follow the best practices of design and development of software systems as the intention is not only to test its performance while following the design principles of an educational social network, but also to be usable for actual integration by educational institutions (UNWE in particular). It is developed on Microsoft .NET Framework 4.5; the open-source edition of ServiceStack [9] is used and extended to provide RESTful hypermedia services; parts of the recently open-sourced Entity Framework [10] are used for building the SQL commands. A compatibility with Mono [11] is continuously verified in order to allow for deployment on Linux machines.

The RESTful API is following the Hydra [12] vocabulary and JSON-LD [13] specification as mechanisms to enforce the user/service interaction graph (also used for predictable caching strategy).

Figure 2: Identifying a node and executing a graph traversal starting from it is initiated by the end-user with this REST-style convention.

The design and development of a specific data model is following the principles of: fluent model configuration; convention over configuration; granularity of required authorization (using claims-based identities – the trusted third party could be a specific implementation for the needs of the EduSN, a university single sign-on system or any other provider like Google, Facebook or Microsoft; they might be treated differently and some of them might be required to use the system or used just as extensions for additional data and endpoints of the user).

The cache (also referred to as in-memory, non-persistent storage) implements a strategy for caching and eviction of data items; their distribution and replication in a multi-server setup; for achieving semantic locality (connected nodes in the social graph being accessible on the same machine). A persistently cached statistics table (updated asynchronously) with data about the determinants of the complexity of the execution of a given operation on a given set of inputs helps with the decisions of which data needs to be cached or evicted.

The implementation of the logical layer is focused around the development of adapters for different types of persistent storages and remote services. A root relational database is introduced to serve as default storage and for persistence of system metadata. An Elgg adapter is developed to provide access to the previous implementation (both for testing and for keeping the access to legacy data). A simple adapter to the Facebook Social Graph API is also developed but is outside the scope of the current testing scenarios.

a. Performance Testing

The purpose of testing the implemented data model is to demonstrate that the user experience criteria can be met; to compare it with the performance of the previous implementation; to identify potential issues and improve both the model and its implementation.

The performance evaluation scenarios are designed to isolate only the parts of the system relevant to the current scope of the experiment.

It includes the testing of the vertical scalability of the model – evaluating the performance on a single server while the dataset is growing in a given pattern and under different types and patterns of workload. While horizontal scalability (evaluating the performance as the data is partitioned and replicated across multiple servers) is an important part of the proposed data model, the testing requires the measurement and analysis of additional metrics like network latency and replication cost and is therefore left as a subject of a separate experiment (it is not comparable to the previous Elgg implementation since it does not support data partitioning directly).

The datasets contain real data from internal UNWE databases. The node degree distribution is generated artificially based on real educational curriculum, courses and groups because it needs to follow power-law distribution (typical for all social network data).

- Dataset A includes 125 professors, 3500 students, 650 groups, 2600 discussions;
- Dataset B includes 250 professors, 7000 students, 1300 groups, 5200 discussions;
- Dataset C includes 500 professors, 14000 students, 2600 groups, 10400 discussions.

The two background workloads used in this test consist of 1) 90% read-heavy and 10% write-heavy operations; 2) 50% read-heavy and 50% write-heavy operations.

During this experiment the following components are isolated and therefore considered insignificant for the interpretation of the results: hardware (running on the same machine, multiple times); relational database engine (running on MySQL, same version, same storage engine, even though other RDBMS might be better suited and provide better overall results); startup slowness (removing the first iterations from the results, a caching warm-up period is usually needed, which leads to slower response times in the beginning).

4. Results and Further Work

The measurements are taken while performing the single specified operation on the behalf of a given user while the workload (as described in the previous section) is running in multiple parallel threads in the background. The dual-layer architecture and the implemented adapters for the physical storages provide an identical interface to any combination of subsystems – we can assume that the tested behavior is not influenced by different higher-level implementations of the selected operations.

Dataset		A		B		C	
Workload		1	2	1	2	1	2
Personalized newsfeed	V1	0.039	0.051	0.057	0.060	0.092	0.099
	V2	0.010	0.014	0.014	0.022	0.020	0.025
User profile read	V1	0.037	0.039	0.044	0.044	0.053	0.058
	V2	0.008	0.020	0.010	0.026	0.011	0.029
Subject group view	V1	0.044	0.050	0.078	0.081	0.102	0.104
	V2	0.013	0.013	0.014	0.015	0.019	0.022
Update status	V1	0.015	0.024	0.019	0.027	0.026	0.030
	V2	0.009	0.011	0.011	0.011	0.016	0.019
Answer a poll	V1	0.018	0.020	0.021	0.022	0.025	0.026
	V2	0.009	0.009	0.010	0.010	0.012	0.014
Post to discussion	V1	0.038	0.039	0.048	0.052	0.065	0.075
	V2	0.015	0.017	0.021	0.023	0.029	0.032

Table 1: Average latency (in seconds) of the six selected operations (three mostly-read and three mostly-write) in the context of the different datasets, workload types, test patterns and intensity. V1 and V2 are the previous (Elgg-based) and the current implementations.

The results can be a subject to further tuning and customization of the databases. However, the results show that the proposed data model serves its very specific and contextual purpose. The underlying causes of the differences in performance are in the nature of the educational social graph – as described in [7], its traversal is much faster when the references to the adjacent nodes are available locally at each step. Putting the proposed dual-layer data model on top of slow (in terms of graph traversal) persistent storages allow the development of customized solutions independently of the inner workings of the database engines.

In terms of testing the vertical scalability of the proposed data model, a LinkBench-based dataset and workload should be used for testing as they match those of Facebook [14] and can be considered appropriate while building a type of a social networking service. It requires a Java-based implementation of the abstract class GraphStore, which needs to communicate with the API of the functional model.

In terms of testing the horizontal scalability of the data model, a multi-server setup is needed along with proper methods for measuring network traffic and the data redundancy caused by the replication (necessary for achieving semantic locality of the graph traversal).

Additional indicators for optimizing the cache contents should be investigated as well – preliminary analysis of the flow of data in the in-memory storage shows that more than 20% of the items that are being selected for caching because of their high execution cost are rarely accessed during the testing. As the real behavior of the users might differ, a manual configuration of the weight of the different persistent storages might be needed as no

statistics is available at the beginning or when new features are being introduced into an already running system.

References

1. 5 Top Social Media Trends For 2014, InformationWeek, 11/2013, (http://www.informationweek.com/software/social/5-top-social-media-trends-for-2014/d/d-id/899805).
2. ResearchGate Wants To Be Facebook For Scientists, Alex Knapp, Forbes, 03/2012, (http://www.forbes.com/sites/alexknapp/2012/03/15/researchgate-wants-to-be-facebook-for-scientists/).
3. Enterprise Social Networks, Fabric Overlays Or Connected Services? , Ben Kepes, Forbes, 01/2014, (http://www.forbes.com/sites/benkepes/2014/01/03/enterprise-social-networks-fabric-overlays-or-connected-services/)
4. Wodzicki K., E. Schwämmlein, J. Moskaliuk, *"'Actually, I Wanted to Learn': Study-related knowledge exchange on social networking sites,"* The Internet and Higher Education, vol. 15, no. 1, pp. 9-14, 2012.
5. Popescu E., D. Cioiu, *"eMUSE - Integrating Web 2.0 Tools in a Social Learning Environment"*, 2011.
6. Vassileva J., *"Toward Social Learning Environments,"* vol. 1, no. 4, pp. 199-214, 2009.
7. Rodriguez M. A., P. Neubauer, *The Graph Traversal Pattern*, chapter in Graph Data Management: Techniques and Applications, 2011.
8. Vicknair C., et. al., *A comparison of a graph database and a relational database: a data provenance perspective*, ACM SE '10 Proceedings of the 48th Annual Southeast Regional Conference, 2010.
9. ServiceStack Library, (https://github.com/ServiceStackV3).
10. Entity Framework Gets Open Source Boost, D. Ramel, 08/2013, (http://visualstudiomagazine.com/blogs/data-driver/2013/08/ef6.aspx)
11. Mono compatibility, (http://www.mono-project.com/Compatibility).
12. Hydra vocabulary, (http://www.w3.org/community/hydra/).
13. JSON-LD, W3C Recommendation, (http://www.w3.org/TR/json-ld/).
14. Armstrong T. G., V. Ponnekanti, D. Borthakur, M. Callaghan, *LinkBench: a Database Benchmark Based on the Facebook Social Graph*, SIGMOD'13, 2013.

Transition from a Generic to a Specialized Implementation of an Educational Social Network

Venko Andonov[1]

[1] UNWE, 1700 Sofia, Bulgaria
vgandonov@gmail.com

Abstract. The standard approach to building an Educational Social Network is using a generic software system (not designed and developed with educational context in mind). Such an implementation at the University of National and World Economy (using the open source product Elgg) has shown that while this is a working solution, multiple issues arise, most significantly - the lack of education-specific features and customizability; and the performance and scalability management. In an attempt to implement the design principles of an Educational Social Network, a new, specialized data model has been developed. In order to transition smoothly to the new model, several migration approaches are evaluated by considering the manageability and the effect on the performance while working with fresh, historical and mixed data. Experiments show that keeping the current system working and integrating its API with the new model is the most feasible solution because of the historical devaluation of the shared content.

Keywords. educational social network, data model, migration.

1. Introduction

An educational social network (Social Learning Network, Social Learning Environment) is a type of online social network which provides a framework and a set of tools and services to enrich the learning process for all of the participants in a given educational environment (in this case, the higher education is the assumed context) [1, 2]. In the case of the University of National and World Economy the open source solution Elgg has been used to provide such an environment (for constrained educational communication and interaction) for the students and the lecturers. The existing implementation is of a generic open source social network system and therefore does not include features relevant to the university academic experience.

A new implementation is being designed and developed to serve the purpose of a specialized educational social network framework - it is private from the university point of view, functionally integrated with internal and external e-learning environments and social networks. Multiple performance bottlenecks were identified in the previous solution and are being addressed in a new model. The tests that were performed (not discussed in the current paper) show improved performance mostly due to a context-aware in-memory graph storage. That suggests a probable good performance on the otherwise slower data model of the existing system – this could be used to perform a migration to the new system without data availability loss, functionality loss and disrupting the users' workflow and overall experience.

2. A Specialized Implementation of an Educational Social Network

The design principles of an Educational Social Network are a collection of requirements for such an environment stemming from existing research and experiments and the shortcomings of the previously implemented and integrated open source solution at the University of National and World Economy. Each principle is described shortly in order to put the new data model in context.

Authenticity of the participants – each profile in an Educational Social Network should correspond to a real participant – student, lecturer or another accepted role. Ensuring the authenticity requires integration with a trustworthy data source of the educational institution. A direct access to internal databases is not recommended due to potential security issues and the tight coupling between the systems. A Single Sign-On approach is preferred where the internal university systems are treated as a trusted authentication provider.

Privacy - keeping the privacy of the participants is a very difficult issue for any social network [3]; it is important to follow the practices of the leaders in the social networking field because the users are familiar with them and there would be less confusion and risk for sharing information with a broader audience than the expected.

Compliance with the organizational policies - if the Educational Social Network is used within an educational institution, it should be compatible with any relevant organizational policies. Compliance with the law (e.g.: in the area of intellectual property), anti-spam measures, detection of offensive behavior. Logging on multiple levels (client, web server, databases, etc.) is recommended both for automatic analysis and manual investigations.

Security – following the best practices in the design and development is a way to protect against application vulnerabilities, data theft and similar concerns. Two-factor authentication is used by many popular websites and if it can be implemented in the specific context is also a layer of protection.

Performance and scalability - the performance of an Educational Social Network is important for the quality of the user experience and for the running costs. A lot of components could cause issues: available hardware; system software; network bandwidth; the database engines; the data model – type and implementation; the application framework; the client scripts. Online social networks are known for having difficulties in scaling up [4].

Multilingual content and interface - it should be useful in all kinds of different countries and cultures – proper localization and internationalization has to be implemented.

Freedom of the teaching styles - the workflow should not be predefined. The lecturer should be free to select or define their own workflow. Restrictions of the organizational policies may constrain this process.

Integrated social environment – the system should be able to build a two-way connection (persistent or on-demand, depending on the technical compatibility) with the popular social network sites and services.

Integrated learning environment - a requirement for building interrelations with different kinds of e-learning systems – both internal (like a Moodle instance) and external ones (like Coursera, edX, etc. [5]).

An environment for the expression of individual knowledge and skills – the service should provide methods for sharing personal information in a desired scope.

A teamwork environment - allow for both synchronous and asynchronous communication between the participants.

Active intelligent environment - actively suggest appropriate people (to connect to) and content (to interact with) to the participants [6].

Distributed and automatic management – distributed content and behavior moderation is preferred. Automatic spam filtering and similar algorithms could avoid several management issues.

3. Data Model and Integration

As shown in Figure 1 the proposed data model consists of two main layers exposing different levels of abstraction to the developer. The logical layer is an abstract graph data structure, providing unified access to the in-memory, persistent and remote storages and services. The educational social graph $G = (V, E, \lambda)$ is a multimodal and multirelational, directed graph with key-value properties attached to each vertex (node) and edge. The physical storage of this graph is decided by adapters, which process a subgraph that has been passed to them together with its metadata in order to retrieve or modify the data.

The functional layer aims to model the specific educational context and workflow of the educational institution. It includes a set of basic data structures and operations that are valid for all social networking services. The specific parts are defined in object-oriented terms.

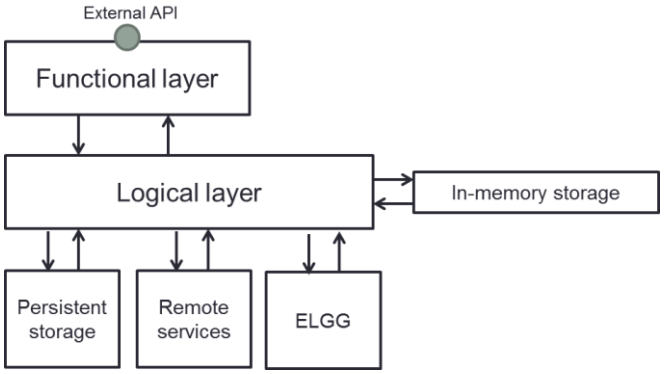


Figure 3: The suggested integration of the new data model and architecture with the existing running ELGG-based Educational Social Network.

The in-memory storage implements a strategy for caching and evicting data items and their distribution and replication in a multi-server setup (for the purposes of horizontal scalaiblity). A persistently cached statistics table gets updated asynchronously with data about the determinants of the complexity of the execution of any given operation on a given set of inputs. This helps with the decisions of which data items need to be cached, evicted or kept available. The implementation of the logical layer is focused around the development of adapters for the different types of persistent storages and remote services. A root relational database is introduced to serve as default storage and for persistence of system metadata. The Elgg adapter for the existing educational social network is developed. The approach is to use

the API provided by the existing system and make it available as a storage adapter to the logical layer.

The migration approaches that are being considered are:

- keeping the existing system intact and developing an adapter in the logical layer to access it through the functional layer API;
- transforming and moving the data to the new persistent storage all at once and shutting down the previous system.

As the second approach would interrupt the users' workflow (in the first scenario there is no need to stop the existing application and GUI, while in the second one this is obligatory) the only issue that would prevent the success of the first approach is if it causes increase in the latency of the operations. The testing scenarios are prepared in such a way so that the performance is evaluated while the user is working with fresh, historical and mixed data.

4. Results and Conclusions

Four operations of the functional model are selected – two read-heavy and two write-heavy. The frequency of the execution of these operations is expected to be high as existing research experiments on educational social networks suggest that these features are among the most used ones by the participants. In the first dataset (fresh data) the operations do not access the previous system in any way; in the second – all accessed and modified data resides on the legacy database; in the third – each operation's input is prepared so that during its execution it reads (and writes) from both persistent data storages.

Operation / Data location mix	Personalized newsfeed	User profile read	Start a discussion	Update a personal status
Completely fresh data access	0.020	0.011	0.029	0.016
Completely historical data access	0.028	0.029	0.036	0.017
Mixed fresh data access and historical data access	0.021	0.026	0.033	0.017

Table 2: Average latency of a read-heavy and a write-heavy operation on datasets with different persistent locations.

The small differences between the latency in the different data access scenarios confirm the suggestion that the proposed new data model manages to provide an appropriate caching layer, agnostic to the issues of the underlying persistent storage engines (the caching strategy depends on the expected and actual behavior of the persistent storages – however, at least in this test setup, these differences do not seem to affect the final result). The historical devaluation of the shared content would reduce the issue even further. Most frequently executed queries in an educational social network are retrieving fresh data as the social graph is very dynamic and most users are interested only in very recent shared content. That is why

it is feasible to trade off very slightly slower query and command execution on historical data for a stable and satisfactory user experience.

References

1. Vassileva J., *"Toward Social Learning Environments,"* vol. 1, no. 4, pp. 199-214, 2009.
2. Wodzicki K., E. Schwämmlein, J. Moskaliuk, *"'Actually, I Wanted to Learn': Study-related knowledge exchange on social networking sites,"* The Internet and Higher Education, vol. 15, no. 1, pp. 9-14, 2012.
3. Gross R., A. Acquisti, *Information Revelation and Privacy in Online Social Networks*, Workshop on Privacy in the Electronic Society (WPES), 2005.
4. Hoff T., *Why are Facebook, Digg, and Twitter so hard to scale?*, High Scalability, 2009, (http://highscalability.com/blog/2009/10/13/why-are-facebook-digg-and-twitter-so-hard-to-scale.html).
5. Lewin T., *Universities Reshaping Education on the Web*, 07/2012, (http://www.nytimes.com/2012/07/17/education/consortium-of-colleges-takes-online-education-to-new-level.html)
6. Varlamis I., *Application of Social Network Metrics to a Trust-Aware Collaborative Model for Generating Personalized User Recommendations*, The Influence of Technology on Social Network Analysis and Mining, Lecture Notes in Social Networks Volume 6, 2013, pp 49-74

Conditions for Implementing Business Process Management Software in Terms of Systems Integration - a University View

Ivan Belev

UNWE, 1700 Sofia, Bulgaria
ivan_belev@yahoo.com

Abstract: The report describes briefly the most common ways for integration between Business Process Management software and existing enterprise systems and data. Basic conditions are highlighted for the organizations in order for them to successfully integrate their IT environment with Business process management software. Data from a research in Universities is analyzed to show if the stated conditions are met or not.

Keywords: Business Process Management software, BPM, systems integration, conditions, University

1. Introduction

The report is focused on Business Process Management technology in terms of systems integration and the area of interest is Universities as organizations that could potentially benefit from implementing a BPM solution. The goal is to try to answer the question – "Do Universities in Bulgaria have the technical means to implement a BPM solution into their IT environment?". For answering this question a few introductory notes have to be made first.

For start it is mandatory to give some information about the main topic that is discussed – Business Process Management software (BPM). A good common definition that applies to most of the BPM software products is:

"Business Process Management software helps to discover, document, automate and accelerate business processes, manage and monitor business performance and enable the organization to continuously optimize and improve its business processes."

The key aspects here are:

- The BPM process starts with discovering and documenting the business processes;
- After that the processes are optimized so that a certain level of acceleration and optimization is achieved;
- When the processes are implemented the software helps the organization to manage and monitor the them in order to discover new ways for further optimization;
- The BPM project doesn't end with the launch of the optimized processes. On the contrary – the organization continues to monitor and remodel the business processes further so that they keep getting better and better.

The next term that is important is systems integration. Looking at it from the perspective of Business Process Management implementation projects the following definition applies:

"Systems integration is bringing together all the necessary software systems and/or data in an organization so that a complete environment can be created for the implementation of Business Process Management technology."

Most of the major BPM vendors provide several common integration mechanisms that can help achieve the required level of integration. Some of them have been used as a base for a survey that was carried out to help achieve the goal of this report. Therefore a short description for each of them is appropriate.

2. Service Oriented Architecture (SOA) Approach

The most common integration scenario in most of the Business Process Management software products is leveraging the Service Oriented Architecture technology by using Web Services. In this scenario the organization should have what is called an ESB – Enterprise Service Bus. The ESB acts as an intermediate communication layer that serves the sole purpose of carrying out the integration tasks in the IT environment. All the different applications make service calls to the ESB, not caring which data source or application will provide or receive the information. The ESB reads or writes the specific information from/to one or more sources and delivers the results.

3. Database Integration Approach

A very common integration scenario will use database integration techniques. There are two aspects of these techniques:

- The more preferable case is using database stored procedures or functions. In this scenario the Business Process Management application will call prebuilt database procedures with the correct number of specific parameters. The procedures will gather or write the data and will return the result. If a database design change occurs – the change can be carried out only on the procedures part and the communication will not break.
- The less preferable case is using database queries to read and write data from specific tables, views etc. This scenario is less preferable as a great deal of maintenance is required with each change in the database design. Also a lot more collaboration is needed in the project implementation process between the BPM team and the Database team.

4. Application Programming Interface (API) Approach

Another common integration scenario is using Application Programming Interfaces (APIs) that an application provides. Although using Web Services is considered using a type of Web API, the approach is somewhat different. The application API provides predefined methods that are prebuilt in the application and can be used to carry out various operations. This may or may not be enough to complete the required level of integration. When using the SOA approach – web services are built specifically for the project based on the requirements of the business process. Using Web APIs is the most common case for using this integration approach, so this was used as a part of the survey.

5. Structured File Formats Approach

The creation and evolution of structured file formats like XML, CSV and others makes it possible for the application to exchange information via data extracts and data imports in previously defined structured file format. This scenario however is not very appropriate since it is not flexible enough.

6. Other Services

In order for the picture to be complete a few other types of integration services may be required:

- Authentication service (LDAP) – required so that the user access to the BPM environment can be easily managed in a centralized repository;
- Mail service – required for carrying out notification tasks and delivering information to different users;
- Message Broker – another intermediate messaging environment that can be used as a communication layer;
- Enterprise Content Management – in some integration scenarios a business process may need to read or write documents to a centralized document management system, so that these documents can be available in other systems as well.

Although each of these cases in important and there are also many others, the most common integration need in this group is considered to be the LDAP provider. Therefore only this case was used as a part of the survey.

7. Survey Questions

Using the described common integration approaches a survey was carried out in several large Universities in Bulgaria. The respondents of the survey were IT personnel from each University that had a complete view of their University's IT environment and were also authorized to provide the information needed. The survey presents only summarized information and doesn't reveal information about the individual units. The following questions were included, grouped by the described integration techniques with available answers – YES and NO:

1. Service Oriented Architecture (SOA) Approach – Web Services:
    a. Do you have in your existing IT infrastructure an Enterprise Service Bus (ESB) that can carry out the communication between different applications?
    b. Do you have in your existing IT infrastructure another system that can be used as a Web Services environment (such as an application server - J2EE, Microsoft IIS, IBM WebSphere Application Server, JBoss, Glassfish or similar)?
    c. Do you have Web Service development or support specialists in your University?
2. Database Integration Approach
    a. Do you have in your existing IT infrastructure structured database systems for storing the application data?
    b. Do you have database administrators in your University?

   c. Do you have database store procedure developers in your University?
   d. Would you allow direct database queries to your databases for the purposes of implementing a new software product?
3. Application Programming Interface (API) Approach
   a. Do you have web applications in your existing IT infrastructure?
   b. Do you have web administrators in your University?
   c. Do you have in your existing IT infrastructure web applications that provide Web Application Programming Interface (API)?
4. Structured File Formats Approach
   a. Do you have in your existing IT infrastructure applications that can import and export data in a structured file format, such as XML and CSV?
5. Authentication service (LDAP)
   a. Do you have in your existing IT infrastructure a single LDAP provider for all existing users (such as Microsoft Active Directory or other)?

Another group of questions was added to check the level of knowledge for the BPM technology and also to point out if the Universities are open to explore such technology:

6. Business Process Management technology
   a. Are you familiar with the BPM technology concept?
   b. Do you have a BPM software product in your existing IT infrastructure?
   c. Do you think that BPM technology can be useful for optimizing the management processes in the University?

8. Survey results

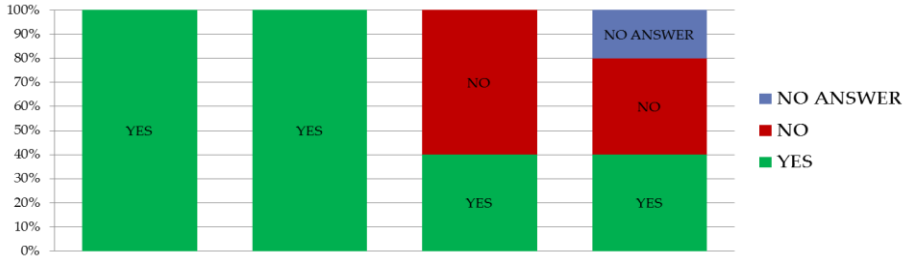
For the purposes of presenting the results the questions have been paraphrased so that the text is shorter. Grouped charts have been used that follow the order of the questions from left to right.

1. Service Oriented Architecture (SOA) Approach – Web Services:

a. Do you have an ESB?
b. Do you have other system to host Web Services (Application server)?
c. Do you have Web Service development or support specialists?

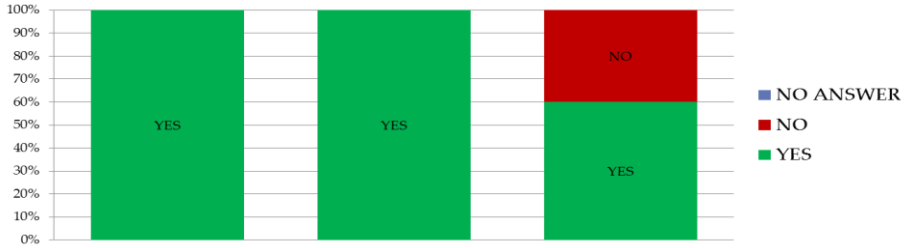
A very small percentage of the Universities have an ESB (only 20 %), but more than a half of them state that they have other means of developing and deploying Web Services.

2. Database Integration Approach
a. Do you have database applications?
b. Do you have database administrators?
c. Do you have stored procedure developers?
d. Would you allow direct database queries to your application databases?



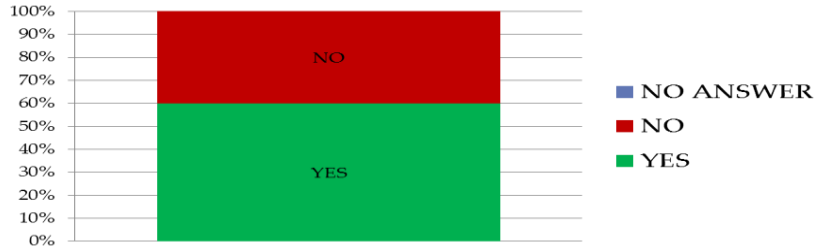
All the Universities that participated in the survey answered that they use and administer their own database applications but most of them lack stored procedure developers. Only 40% of the respondents answered that they would allow direct queries as a database integration technique. Some of the respondents chose not to answer the question which may be due to security considerations or lacking the authority to answer such a question.

3. Application Programming Interface (API) Approach
a. Do you have web applications?
b. Do you have web administrators?
c. Do you have web applications that have some form of Web APIs?



All of the respondents claim they have web applications and web administrators. Most of the Universities that participated employ some form of APIs in their web applications.

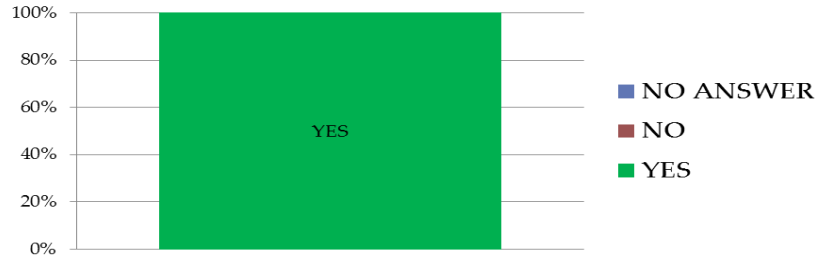
4. Structured File Formats Approach
a. Do you have applications that can input and output data in a structured file format such as XML?



Surprisingly not so many of the respondents answered yes at this question.

5. Authentication service (LDAP)

a. Do you have a single LDAP source for all users in the organization?



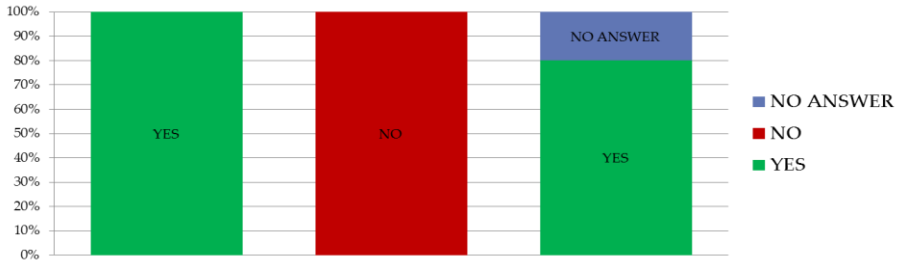
All the respondents answered this question with YES which was somewhat expected.

6. Business Process Management technology

a. Are you familiar with the BPM Software concept?

b. Do you have BPM software?

c. Do you think that BPM Software will be useful in your University?



The answers to last group of questions were not surprising at all. Again there were empty answers on one of the questions.

9. Conclusion

Now that the results are present the following conclusions can be made based on the provided answers:

- The results from the survey show that only a few of the conditions for integration of Business Process Management Software with existing software systems and data are consistently met in the Universities that participated in the research.

- This points that most of the organizations will have difficulties implementing BPM software into their environment.

References

1. IBM's Business Process Management Book Second Edition, 2013, (www.ibm.com)
2. Oracle BPM Suite Integration Features, (www.oracle.com)
3. Kolban's Book on IBM Business Process Management, September 2013 Edition
4. BPM technology review 2013: Assessment Framework, June 2013, (http://www.mwdadvisors.com/library/detail.php?id=521)
5. BPM technology: Appian June 2013, (http://www.mwdadvisors.com/library/detail.php?id=519)
6. BPM technology: Pegasystems June 2013, (http://www.mwdadvisors.com/library/detail.php?id=516)
7. BPM technology: Software AG June 2013, (http://www.mwdadvisors.com/library/detail.php?id=515)
8. BPM technology: TIBCO June 2013, (http://www.mwdadvisors.com/library/detail.php?id=514)
9. Survey among several Universities in Bulgaria conducted by the author for the purpose of the report.

IT Tendencies in Banking Systems Security

Lili Emiliyanova Kitova

lili.kitova@gmail.com

Abstract. Data security and cyber security are always important issues for financial services firms. After all, almost every business line in the industry depends on good, clean data. Until recently, however, cyber security hadn't received all that much attention, since firms had traditionally done a pretty good job of protecting against attacks.The Internet has played a key role in changing how we interact with other people and how we do business today. As a result of the Internet, electronic commerce has emerged, allowing businesses to more effectively interact with their customers and other corporations inside and outside their industries. One industry that is using this new communication channel to reach its customers is the banking industry. And in 2012 and 2013, bank technology budgets continue to increase. But with the economic recovery still on shaky legs, and with regulatory scrutiny more intense than ever, banks' IT investments are likely to be focused largely on driving efficiencies and complying with new requirements. Bank Systems & Technology identifies the IT trends and hot technologies that will change the game in the year ahead. IT tendencies in banking system security are related with Electronic Banking, Mobile Banking Security, Information Security Trends, Online banking systems security.

Keywords: Mobile Banking Security, Information Security Trends, Online banking systems security, Electronic Banking

1.Introduction

Trying to stay ahead of the curve when it comes to IT issues is a challenging task. Emerging technology forces in the financial services industry are already impacting business. The convergence of these forces does present challenges; however, it also provides a window of opportunity for financial institutions to elevate business performance and gain a competitive advantage. Proficient provides a monthly perspective on some of the most talked about IT issues and emerging trends to help industry professionals identify and rationalize their IT investments.

For the past several years, bank IT budgets generally remained flat. The financial crisis and ensuing fallout forced belt-tightening across the industry. In 2011, however, bank IT executives finally enjoyed some breathing room thanks to some revitalized spending power.

Mobile banking started as a novelty, something only techies and first adopters felt comfortable using. But as smartphones have skyrocketed in popularity over the past few years, mobile banking adoption has increased along with it.

Initially, many banks' mobile offerings consisted of their online banking model ported to an iPhone or Android device. As mobile has grown into a maturing channel, however, banks and their vendor partners have produced richer mobile offerings that take advantage of its unique capabilities. And the rise of the tablet gives financial institutions another unique interface through which to interact with consumers.

2.IT tendencies in banking systems security

The tactics used by cyber-criminals to target sensitive financial data are sophisticated and constantly changing. So, too, must the security controls financial institutions have in place to stop the next cyber-threat. But as technological innovation brings promise and excitement to the financial services industry, it also brings new opportunities for fraudsters and hackers. Here are five trends we can expect to see in the future of cybersecurity.

Dr. David Chaum, CEO of DigiCash said that "Security is simply the protection of interests. People want to protect their own money and banks their own exposure. The role of government is to maintain the integrity of and confidence in the whole system. With electronic cash, just as with paper cash today, it will be the responsibility of government to protect against systemic risk. This is a serious role that cannot be left to the micro-economic interests of commercial organizations." The security of information may be one of the biggest concerns to the Internet users. For electronic banking users who most likely connect to the Internet via dial-up modem, is faced with a smaller risk of someone breaking into their computers. Only organizations such as banks with dedicated Internet connections face the risk of someone from the Internet gaining unauthorized access to their computer or network. However, the electronic banking system users still face the security risks with unauthorized access into their banking accounts. Moreover, the electronic banking system users also concern about non-repudiability which requires a reliable identification of both the sender and the receiver of on-line transactions. Non-secure electronic transaction can be altered to change the apparent sender. Therefore, it is extremely important to build in non-repudiability which means that the identity of both the sender and the receiver can be attested to by a trusted third party who holds the identitycertificates.

2.1. Risk-Based Authentication in the Cloud.

Financial institutions find it much easier to respond against known malware attacks than they do against unknown attack vectors or zero-day vulnerabilities. That's why one promising approach to security for the industry is the concept of outsourced protection. We may see ID and fraud management being outsourced or moved into the cloud. The cloud providers can provide you with a lot more background intelligence about attacks and issues that may not have hit you, but have hit other people.

By working with multiple institutions on a real-time basis, cloud security providers can muster a more comprehensive defense. They'll be amassing information as to what it means to be normal from a user or a peer group perspective, and then alerting on deviations from that normalcy. That's probably going to be another layer of defense.

2.2. Biometric-Powered Bank Applications.

The big problem with passwords is that they're difficult to remember and easy to store in an unprotected area. Even if an application goes to extreme lengths to avoid storing usernames and passwords within its protected data area, it's hard to stop users from pasting their passwords into an unencrypted notebook page or draft email for quick reference.

Biometrics promises an authentication technique that's easier than remembering (or copying and pasting) a password. One approach to biometrics is voiceprint ID, in which the user is asked to repeat a phrase or a series of digits. The phrase might be the user's home phone number or mobile number; or, to eliminate the "replay attack" risk that an attacker has recorded the real user's voice, the requested phrase could also be a random series of digits, or one of several random phrases.

One of the potential drawbacks of voiceprint IDs is that the user may not be in a quiet area conducive to providing a clean sound sample. In such situations, another promising technique is facial recognition.

Facial recognition, as with voiceprint IDs, might also be vulnerable to "replay attacks" with the current level of the technology.

To counteract that attack vector, one idea is to register facial biometrics as a movie. You'd rotate your head to the left, and then do a 180-degree rotation to the right. At registration, the system would build a three-dimensional model based on that. Then, when you authenticate, you can compare two 3-D images of a head instead of just two 2-D images.

Biometric approaches work best in defined niches. For mobile devices, the quality is still suspect, it's tricky to do, and it potentially compromises privacy. If you do biometrics right, that's great, but it could be more trouble than it's worth.

2.3. Credit Cards with Token Generators.

One of the main problems with token-generating devices is that they're bulky and unwieldy. But what if you could get your one-time password from the credit card in your wallet?

An interesting contender for the out-of-band authentication challenge is having a token generator embedded in the form factor of an ordinary credit card. You can integrate a PIN-protected one-time password circuit on the same device, along with the EMV chip and magnetic stripe.You can use it to swipe, or at an ATM, or to read off a one-time password.

It's the same size as a credit card, and the batteries last about two years. If deployed at scale, the production costs could be reduced significantly while offering strong protection against fraud.

2.4. Polling the Device.

The raw technologies involved with mobile computing give it certain advantages when it comes to authentication. In some ways, mobile is better than a PC for authentication, because you can factor in things such as location. Combinations of transactions can get flagged very quickly, and then blocked and challenged. For example, if successive mobile logins were to occur thousands of miles apart, using verifiable location data based on cell phone towers, the bank may infer that one of the two logins has to be a fake.

Ordinarily, downloaded apps can be granted permission to access the location data associated with a device. Beyond location, there's a big opportunity for banks to strengthen their authentication practices by tapping into the full complement of data and services available on mobile devices.

For example, suppose a bank makes a copy of a user's contact database and then, before authorizing a transaction, checks to see whether those same contacts are present on someone's phone. That way, someone who just steals the user ID and password for a bank account would be unable to log in; the thief would have to steal the person's actual phone, or somehow copy that person's entire address book along with the bank credentials.

Another approach is examining the physical behavior of how people use their smartphones. When looking up phone numbers, some people always search the contacts list, while others type the number first. There are several ways to do the same things, and if you observe someone's behavior over a long period of time, you're going to see repetitive patterns that are different, person by person.

You can argue that this is biometric information, and there would be a huge set of concerns around privacy. Whether it would ever be possible to realize some of these enhanced OS-based techniques in Apple iOS remains an open question. Nevertheless, the

possibilities afforded through digging deeply into the phone's data store are sure to make such ideas hard to rule out entirely.

2.5. Device-Based Authentication.

If security at the device level becomes enough of a differentiator in the market, we may see the industry shift to entirely new business models that place device manufacturers and network operators in the driver's seat.

Mobile operators and device manufacturers were caught flat-footed with the rapid success of Apple iOS, and they'd surely relish the opportunity to figure out some new way to differentiate themselves with a more advantageous bargaining position relative to the operating system companies.

Out of the possibilities involved with having the secure area on a smartphone available— for an annual fee — to financial institutions and other payments, e-wallets, and loyalty application providers. The network operators want to be in charge of that. Part of the delay of NFC coming to market is who controls the secure market, who's paying for it, and who makes revenue from it. Device-based authentication could be embedded into the handset by the device manufacturer or network operator; or located on a removable SIM card or microSD card to be provided by a bank or other player.

With these considerations in mind, banks should keep a sharp lookout and maintain a nimble footing when evaluating the evolving possibilities in the mobile ecosystems emerging within their respective geographic markets.

3. Conclusion

We will see banks continue to develop solutions for multiple channels but using a single set of technology to do so. A cohesive set of technologies will make mobile app and online development easier for banks to manage.

Several high-profile acquisitions, such as Brookfield, Wis.-based core banking software provider Fiserv's March 2011 acquisition of mobile banking and payments software provider Mobile Commerce (M-Com; Atlanta), as signs that the industry is moving in this direction. The reason behind these kinds of acquisitions is because need these tech assets that will allow them to serve their customers.

As banks continue to search for efficiencies and consolidate operations, the convergence of mobile and online technologies looks to be a prime opportunity to do so in 2014.

The Internet has grown exponentially, with more than 30 million users worldwide currently. The Internet enhances the interaction between two businesses as well as between individuals and businesses. As a result of the growth of the Internet, electronic commerce has emerged and offered tremendous market potential for today's businesses. One industry that benefits from this new communication channel is the banking industry. Electronic banking is offering its customers with a wide range of services: Customers are able to interact with their banking accounts as well as make financial transactions from virtually anywhere without time restrictions. Electronic Banking is offered by many banking institutions due to pressures from competitions. To add further convenience to the customers, many banking institutions are working together to form an integrated system such as the Integration Financial Network and the Gendex Bank International.

On the other hand, this has not been readily accepted by its users due to the concerns raised by various groups, especially in the areas of security and privacy. Moreover, there are many potential problems associate with this young industry due to imperfection of the security methods. In order to reduce the potential vulnerabilities regarding to the security,

many vendors have developed various solutions in both software-based and hardware-based systems. Generally speaking, software-based solutions are more common because they are easier to distribute and are less expensive. In order for electronic banking to continue to grow, the security and the privacy aspects need to be improved. With the security and privacy issues resolved, the future of electronic banking can be very prosperous. The future of electronic banking will be a system where users are able to interact with their banks "worry-free" and banks are operated under one common standard.

References

1. http://www.banktech.com/management-strategies/8-bank-technology-trends-that-will-shape/232300804?pgno=1
2. http://www.informationweek.com/security/risk-management/7-top-information-security-trends-for-2013/d/d-id/1107955?
3. http://www.wallstreetandtech.com/data-security/10-financial-services-cyber-security-tre/240143809
4. http://bankinnovation.net/2013/05/top-5-technology-trends-in-financial-services-april-2013/

Methodologies (Methods) for Business Intelligence (BI) Implementation

Ivan Georgiev Andonov

University of National and World Economy (UNWE)
ivanandonov87bg@abv.bg

Abstract: The goal of this paper is to explore and analyze some of the methods for BI implementation. As a result of the analysis, it will be made a comparison of the similarities and differences between the mentioned methods for BI implementation in the paper.

Keywords: Business intelligence, Methodologies for BI implementation

1. Definition of Business Intelligence

Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance. [1]

2. Methodologies of the leading companies in the field of Business intelligence

In this paper, we are going to explore and analyze methodologies for business intelligence implementation of the leading business intelligence tools in this area such as:

- Hewlett Packard;
- IBM;
- Reliable Software BI Project life cycle methodology for implementation of MicroStrategy BI tool.
- Business Intelligence Roadmap

These methodologies are chosen as a result of the analysis of the Gartner group in the area of Business Intelligence Systems. In figure 1 it is shown the magic quadrant of Gartner group with the results as of February, 2013. The mentioned methodologies are applied in the implementation of leading business intelligence tools of IBM, Qlik Tech, and MicroStrategy such as MicroStrategy, Qlikview, and IBM Cognos Business Intelligence.

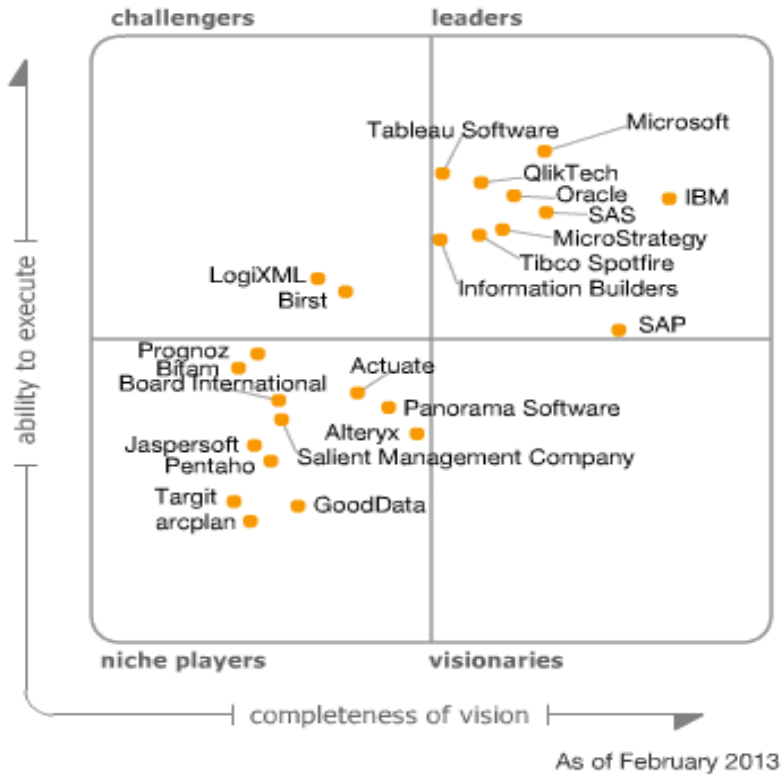


Figure 1 Gartner magic quadrant [2]

2.1. HP global method for business intelligence implementation [3]

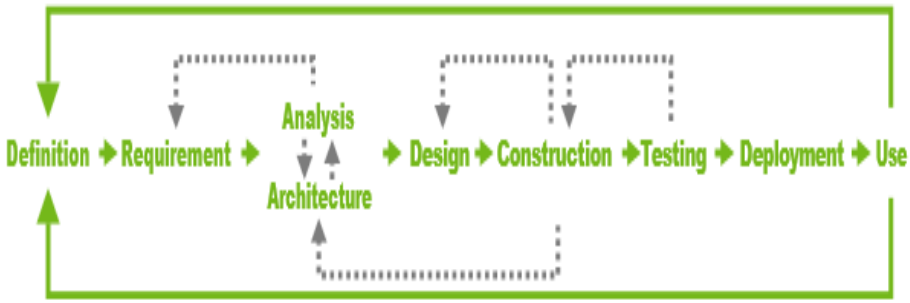


Figure 2 HP global method for business intelligence implementation

This methodology consists of 9 phases – Definition, Requirement, Architecture, Analysis, Design, Construction, Testing, Deployment, and Use.

In **Definition phase** the scope of the project is clearly defined. Here it is important to be conducted preliminary investigation and start-up – to be defined business requirements, technical landscape, considerations such as project management, risk management, project approach.

In **Requirements phase** the goal is requirements to be achievable and accurate, to be captured business and data requirements, technical requirements, and metadata requirements.

In **Architecture phase** it should be created a solution architecture that provides compatibility of the domain architectures,

In **Analysis phase** it is essential what business expects from the new deployed technology. How this business intelligence system will contribute to be increased the value added of the company.

In **Design phase** the most important thing is to be produced documents with sufficient information that during the next phase Construction more developers – junior and senior to participate in the coding.

In **Construction phase** the goal is to be developed and tested a real model of business intelligence solution for the relevant company.

The next phase is **Testing phase** where are conducted Integration Test, System Test, and User Acceptance Test.

In **Deployment phase** the end user training is conducted and the BI solution is in the hands of the business users.

In **Use phase** there are a lot of activities such as monitoring, support of the end users and explore how the business intelligence system could be improved.

2.2. IBM method for Business intelligence implementation [4]

This method (methodology) contains of the stages shown in figure 3 below. These stages are: Analyze Design, Configure and Build, Deploy, and Operate.

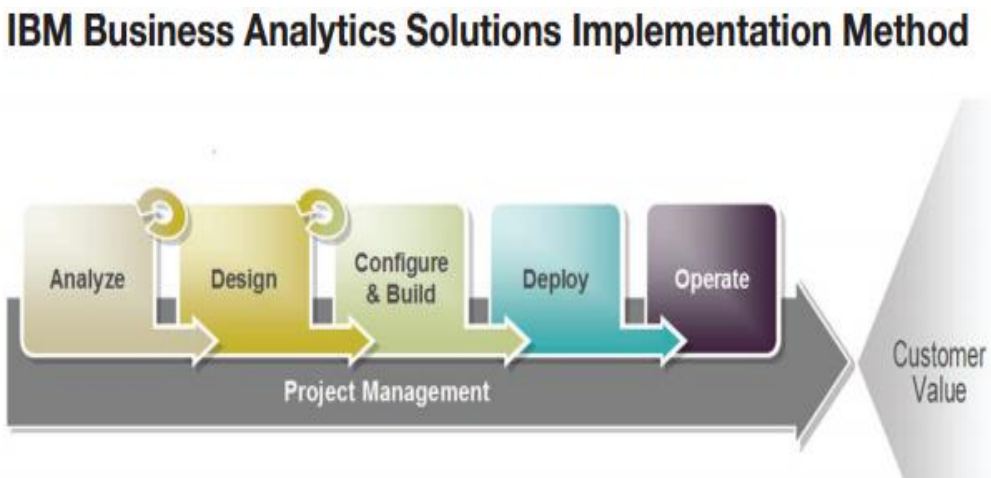


Figure 3 IBM implementation method

The IBM implementation method tries to describe step by step what should be done during the implementation process. This method describes key roles, important requirements

to be clearly defined and everyone in the project team should know his/her responsibilities, deadlines and etc.

Analyze – during this stage it should be specified what the situation in the company is, and what the situation should be when the business intelligence solution is in use.

In **Design stage** all dependencies and components should be defined. During **Configure and Build stage** all component are configured, built and integrated.

In **Deploy stage** plan to run and maintain the solution is prepared. In it there is support schedule.

In **Operate stage** maintenance tasks and checkpoints after roll out are included.

2.3. Reliable Software BI Project life cycle methodology for implementation of MicroStrategy BI tool

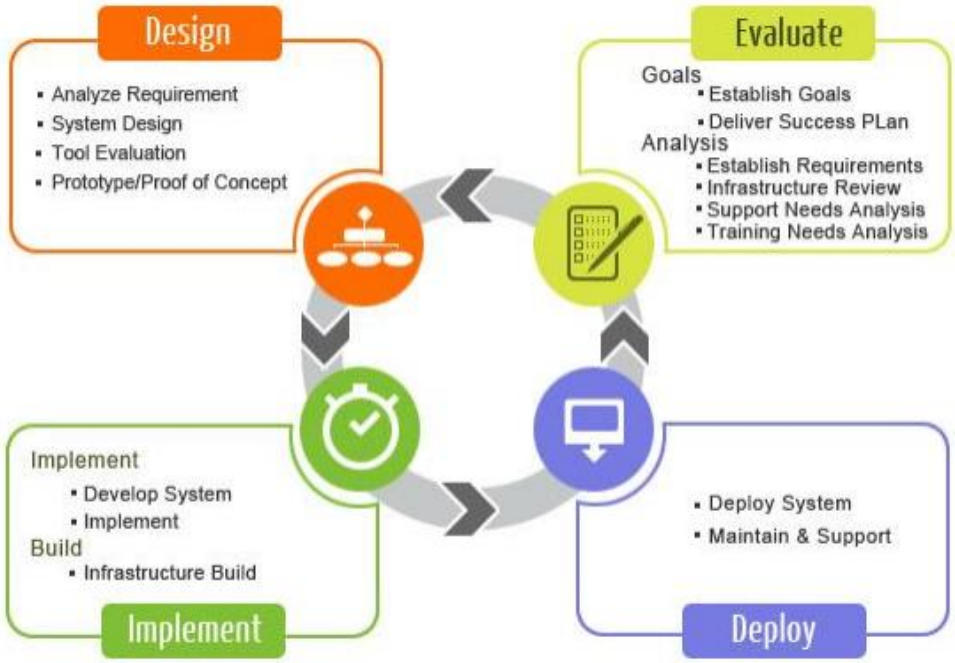


Figure 4 Reliable Software BI Project life cycle methodology

This methodology contains 4 stages – Evaluate, Design, Implement and Deploy [5].

During the first stage **Evaluate** a gap analysis is performed in order to produce a document that provides the required information for the reports that will be generated by the business intelligence tool. Key part of this stage is preparation of plan how the MicroStrategy BI implementation will be done.

On the next stage **Design** tables, reports and all other documents that should be architected in MicroStrategy Project are analyzed and all decisions about it are taken.

During the third phase the system is developed and implemented and the infrastructure is built and the final stage is for validation of the data, final test and support.

2.4. Business intelligence roadmap by Larissa Moss and Shaku Atre [6]

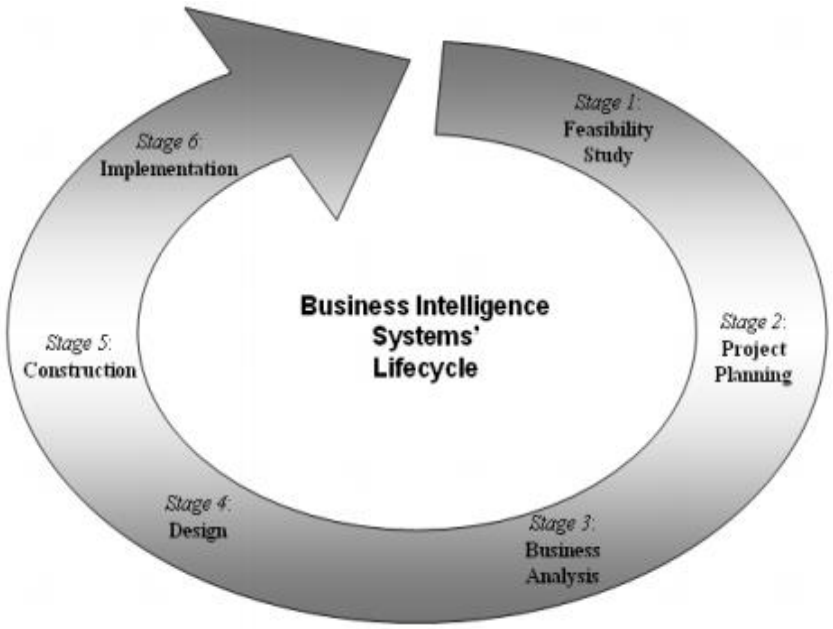


Figure 5 Business Intelligence roadmap

This roadmap for Business intelligence implementation is divided in 6 stages and 16 steps where in details every aspect of the implementation is explained and everything is considered.

During the first stage Business case assessment is made. It is decided whether or not the project to be done – it means whether the return on investment (ROI) is positive or negative. After that, project planning is coming where the infrastructure is evaluated and the project plan is described and key roles are defined.

During the third stage project requirements are defined, Meta Data Repository and Data are analyzed and Application prototyping is prepared.

In Design stage, the design of ETL, Database and Meta Data Repository is completed and should be realized and developed.

During the Construction phase of the implementation ETL, Application and Meta Data Repository are developed and Data mining tool is applicable, data mining enables statistical models to be applied and to be found new knowledge in data. Data mining briefly can be described as it follows: state of business problem; collect, consolidate and cleanse data; prepare data; build analytical data model; interpret data mining results; perform external validation of results; and monitor analytical data model over time.

Conclusion:

To sum up, the mentioned methodologies in this paper are offering stages that have many things in common and the decision which method will be chosen or which methodology, it does not guarantee the success of the implementation of the Business intelligence system(s).

The situation is more complicated – the main resource that is really important is the project team that is in charge of implementation of BI. Every member of the team should be well qualified, experienced and to know what are his/her duties and precisely to do them.

References:

1. Gartner group, available at http://www.gartner.com/it-glossary/business-intelligence-bi/
2. http://www.microstrategy.com/about-us/analyst-reviews/gartner-magic-quadrant
3. HP Global Method for Business Intelligence Implementation
4. IBM Business Analytics Solutions Implementation Method (BASIM)
5. Business Intelligence Tool Migration
6. Business Intelligence Roadmap – The complete project lifecycle for decision - support applications, Addison – Wesley, 2004.

Computer-assisted Courses

Zahari Goranov
UNWE, 1700 Sofia, Bulgaria
zgoranov27@e-dnrs

Absract. Use of Modeling and Simulation in computer assisted trainings for improving civil security becomes particularly relevant to the development of today's society and the introduction of new technologies in everyday life. Contemporary threats of terrorism and natural disasters require coordination of many different structures. Simulation is a way in which we can advance to check what the role should be run.

Keywords. Simulation, computer – assisted course, training

Introduction

In the EU and NATO, computer aided teaching is seen as a powerful tool for learning , analysis and evaluation tool with a scientific contribution to change management in the security sector, the basis for a parallel implementation of a number of technology demonstrations and experiments. In recent years, Bulgaria has started to gain experience with the use of computer simulations.

Furthermore, especially in the real environment risk of casualties is lower - because it provides how to avoid certain risk factors.

Computer assisted exercises help officers to conduct several courses of action on the battlefield and then especially in a real fight to know which one is most effective. On the other side is cheap and saves the taxpayer`s money.

- simulations are - cheaper and safer way to conduct experiments
- simulations can be more realistic than the experiments
- simulations can be performed in a different time
- simulations allow the construction of a coherent medium for analysis of integrated systems simulated
- simulations allow to live in the future

1. Definition

Definition for Computer-assisted courses synthetic doctrine which resources are generated perform actions and run in a simulation environment. The use of modeling and simulation in computer assisted exercises improving civil security becomes particularly relevant to the development of today's society and the introduction of new technologies in everyday life. This allows using the latest technology to recreate the objective reality and the involvement of people in the newly created virtual reality for training purposes. On the other hand the occurrence of events such as global terrorism pose new challenges to the security of citizens and the recreation of virtual situations the security context [1].

Another determination for simulation is:

A computer simulation is a simulation, run on a single computer, or a network of computers, to reproduce behavior of a system. The simulation uses an abstract model (a computer model, or a computational model) to simulate the system. Computer simulations have become a useful part of mathematical modeling of many natural systems in physics (computational physics), astrophysics, chemistry and biology, human systems in economics, psychology, social science, and engineering. Simulation of a system is represented as the running of the system's model. It can be used to explore and gain new insights into new technology and to estimate the performance of systems too complex for analytical solutions [2].

2. Benefits of computer – assisted course

The emergence of new threats to the security of the citizens related to the potential dangers of terrorism requires close interaction between all the elements integrated security sector. The issue of this integration is discussed in detail in and here we note that a scientifically valid methods for integration security sector is the need of recreating imaginary emergencies action, based on the application of the idea of modeling and simulations in computer assisted exercises (CAX) [3].

Superiorities of computer-assisted courses:

- simulation of many variations and situations
- creates a low-cost medium for communication and interaction
- opportunity to simulate the missing element or a participant in the teaching
- modeling of dangerous natural or anthropogenic processes and virtual situations
- create conditions for the measurement of group work or analysis
- possibility of documenting and analyzing the experiment, lessons learned for future events, creating knowledge bases
- experimentation of concepts, software for management support

Economical aspect of the simulation

In these times of austerity, each euro, dollar or pound sterling counts. Smart defence is a new way of thinking about generating the modern defence capabilities the Alliance needs for the coming decade and beyond. It is a renewed culture of cooperation that encourages Allies to cooperate in developing, acquiring and maintaining military capabilities to undertake the Alliance's essential core tasks agreed in the new NATO strategic concept. That means pooling and sharing capabilities, setting priorities and coordinating efforts better. From 2008 the world economy has been facing its worst period since the end of the Second World War. Governments are applying budgetary restrictions to tackle this serious recession, which is having a considerable effect on defence spending [4].

The Republic of Bulgaria has hosted the sixth exercise in the South-Eastern Europe Simulation Network series SEESIM. The SEESIM 12 computer assisted exercise is being conducted in the National Military Training Complex "Charalitza" 14-19 October 2012 and distributed to the following ten remote sites: Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Italy, Montenegro, Serbia, Macedonia, Ukraine, and SEEBRIG HQ in Larissa, Greece. The SEESIM series of civil-military disaster relief exercises has been conducted since 2000 as an initiative of the Southeastern Europe Defence Ministerial (SEDM) process. Each SEESIM Exercise has a two-year planning cycle.

The goal of SEESIM 12 was to establish cooperative relations and common procedures among the countries from the region to facilitate and accelerate the process of requesting and providing assistance in response to natural disasters and terrorist attacks. This

is the first SEESIM Exercise to utilise tools provided by NATO, including computer instruments and models for planning and conducting exercises of this type.

The broad scenario of the Exercise is based on a series of cyber attacks against the governments of the countries in Southeastern Europe followed by biological or chemical terrorist attacks and a large natural disaster. These events result in a breakdown of communications and great loss of life. The Exercise will activate SEEBRIG's Engineer Task Force and the national crisis response centers of participating countries [5]. To use software is much cheaper than the creation of a real situation. Ourthermore may be made a number of situations, to be trained.

Another good example for computer assisted course is Phoenix 2010. Establish a National Staff who plan, organize and coordinate the actions of public authorities to tackle a crisis. The tasks of the "Phoenix 2010" are to improve the standard operating procedures for emergency situations, stabilization and recovery at national level and at international interaction. To achieve the tasks will be used computer simulation systems and models. This is the first such simulation implemented nationwide. The script includes progressively deteriorating situation resulting from a wave of refugees, natural disaster and terrorist act. On the first day of the exercise test will undergo the skills of state institutions to coordinate their actions in a situation of refugee wave [6].

Computer assisted exercises are powerful , integrated and convenient tool for training both the military as well civilian experts in the security sector, as they allow training with high flexibility and low cost [7].

„Bulgarian army relies increasingly on computer-assisted exercises, which are the most best tool for increased efficiency in preparation for achieving full interoperability at low cost and reduce the level of risk compared with actual military exercises" [8].

Reform military education system and the new tasks of the army to impose looking for new and more effective methods of training. In response to these requirements in the process of preparing to apply modern methods such as modeling and simulation. The advantages of the use of modeling and simulation in various aspects, but the most Generally, they are limited to the ability of personnel training and staffs to be carried out in a synthetic medium without exiting subdivision the area, to repeat and training activities in operational planning, operations and actions, as well as to analyze and learn from practice.

The main objectives related to the use of systems modeling and simulations are: - Integrating and sharing capabilities of the built computer-assisted exercises (CAX) and simulations and creating technological and technical conditions to improve operational the headquarters and units of the Bulgarian Army.

- Providing information and communication environment conducting joint distributed CAX with staffs and formations NATO. Conducting experiments , forecasts and analyzes the process of operational planning .

- Creating conditions for development in the Republic of Bulgaria regional center for computer assisted exercises of countries Southeastern Europe. Development of integrated modeling and simulation Republic of Bulgaria as an open system in accordance with the development of NATO capabilities in the application of models and simulations in the interest of defense, in particular the creation ofNetwork of NATO training and education (NATO Education and Training Network).

• Joint Conflict And Tactical Simulation (JCATS) - Implemented in NVUK " Charalitsa " BA " D. GSRakovski " NMU " V. Levski " Naval Academy " N.Vaptzarov . " Established local networks in these simulations centers integrated into the global network of the Ministry of Defence and Bulgarian Army. The functional scope and operational- tactical

JCATS capability is enhanced by the addition of software tools to manage events and incidents (Joint ExerciseManagement Module) and scenario (Joint Exercise Scenario Tool) and other compatible tools developed under the programs

NATO in the field of simulation.

• System Simulation training of formations Bulgarian army with virtual simulation model (Virtual Battle Space (VBS2)). Scope of delivery are computer systems, components build a local network operating system and specialized Software - VBS2 Virtual Training Kit (VBS2 VTK). The software works on the operating system Windows. The product contains the VBS2 VTK virtual environment with scenario editor module for generating terrain and three-dimensional models of objects and integrated entrance (gateway) to other simulation systems based on standards High Level Architecture / Distributed Interactive Simulation (HLA / DIS). The software package 30 includes tools for analysis and statistical treatment of results.

It consists of the following software components:

- Simulation module "Desktop Trainer";
- Module Development "Development Suite";
- HLA / DIS Gateway (LVC Game).

The module "Desktop Trainer" provides an opportunity to generate and visualizing synthetic medium containing the elements of the terrain, infrastructure, equipment and personnel. The module contains the editor scripts, a library of training scenarios and a library of three-dimensional (3D). Standard models of objects - military equipment and armament (consideringtactical and technical characteristics), equipment and personnel. Perform be simulated training scenarios at a tactical level (group and individually), rehearsing missions and joint operations.The module "Development Suite" provides tools for creating and in the script of new sites - special characters , weapons systems ,equipment elements of the terrain , infrastructure.(9)

I will give another example for simulation and education (and how to save money).

Flying an F-16 fighter costs an estimated $5,000 an hour, compared to $500 per hour in a simulator. Driving a tank costs $75 per mile; a tank driver simulator, $2.50 per mile. Operating an Apache helicopter cost $3,101 per hour; a simulator, $70 per hour. Sometimes, existing computer software can be bought off-the-shelf and modified to provide military training at a fraction of the cost of traditional methods. Ensign Herb Lacy, a 24-year-old preflight student at the Pensacola Naval Air Station, recently bought a copy of a Microsoft game called Flight Simulator for about $50. He then modified the program to recreate the appearance and controls of a T-34C Turbo Mentor, which the Navy uses for primary flight training. The total cost was less than $250. Lacy logged 50 hours of flight time on his personal computer before he ever climbed into a real T-34C.

The Navy was so impressed by Lacy's accomplishment that it has begun operating six makeshift T-34C simulators, using the same software, at a cost of $6,000 each, compared with millions of dollars for conventional simulators. Another attraction of simulation is that it is safer than traditional forms of military training, officials said. Military training, they noted, is inherently dangerous. In fact, they pointed out, the only U.S. casualties of the air war in Yugoslavia were two Army pilots who were killed when their Apache helicopter went down on a training mission in Albania. The increasing use of simulation devices is playing a significant role in improving safety in military aviation, according to NTSA. The accident rate for naval aviation has fallen from 20 per 100,000 hours in the 1950s to 2.39 per 100,000 at most recent count. Even with such accomplishments, simulation will never "replace live training entirely," cautions Air Force Col. Paul Avella, chief of conventional operations and

training for the Air Combat Command, Langley Air Force Base, Va. Instead, he predicts, "it will enhance and expand traditional training." [10].

3. Conclusion

In conclusion we can say that computer modeling will take an increasingly important both in the military and in civilian life. Modeling exercises and the use of simulators for training is much cheaper than a genuine workout. In the future we should continue to use the computer as a learning and training tool.

References

1. Brigade General Dimitar Dimitrov, 2009
2. Mintchev,Z Use of modeling and simulation in computer-aidedexercises to improve civil security
3. Shalamanov.Vq Tomov.N, Minchev.Z, Nicolaeva.I 2005
4. http://en.wikipedia.org/wiki/Computer_simulation
5. http://www.nato.int/cps/en/natolive/topics_84268.htm?
6. http://seesim.mod.bg/index.php?id=1&nid=124
7. http://phoenix.mod.bg/index.php?id=43&nid=50
8. http://www.mod.bg/bg/EXT/InstitutOtbrana/docs/Concept_Strategy_S&T.pd
9. http://www.nationaldefensemagazine.org/archive/1999/November /Pages/Simulation4398.asp.

Evolution of the European Grid Infrastructure from Grid to Cloud

Vladimir Dimitrov

Faculty of Mathematics and Informatics, University of Sofia,
5 James Bourchier Blvd., 1164 Sofia, Bulgaria
cht@fmi.uni-sofia.bg

Absract. The computational ¬¬Grid is an independent, comprehensive, overall hardware and software infrastructure, which offers inexpensive access to high volume of computational resources. Realization of Grid is the European Grid Infrastructure (EGI). The EGI is gradually evolved and matured as a result of the projects DataGrid, WLCG, EGEE and EGI-InSPIRE. The beginning of the EGI infrastructure was placed in 2001, when the foundations of the infrastructure were built and the first middleware packages were developed. During these projects the middleware software packages evolved through: EDG middleware, LCG middleware, gLite middleware to EMI middleware. For the last two grid environments – gLite and EMI, some service-oriented approaches were considered. Unfortunately, the idea for the service-orientation was partially accomplished. In 2013 the EMI project ends and currently there is not new project specially dedicated for grid middleware development. This was defined as a milestone for EGI and grid computing in Europe and brings the evolution of the EGI infrastructure to the new turning point. Will cloud computing influence the evolution of the infrastructure and how, all these problems will be considered in the paper.

Keywords. European Grid Infrastructure, service-oriented architecture, business processes, Cloud computing

1. History of the European Grid Infrastructure

The computational Grid [1] is an independent, comprehensive, overall hardware and software infrastructure, which offers inexpensive access to high volume of computational resources. Realization of Grid is the European Grid Infrastructure (EGI) [2]. The EGI is gradually evolved as a result of many projects. The beginning of the infrastructure was placed in 2001, when the European DataGrid (EDG) [3] project was launched. The aim of the project was to develop grid middleware for the needs of the scientific European society, which to allow sharing of data and computational resources.

Parallel with the EDG project the Worldwide LHC Computing Grid (WLCG) [4] project was started. The project was launched in 2002. The purpose of the project was to prepare computational infrastructure for simulation and analysis of the data generated by the Large Hadron Collider (LHC) at CERN. The newborn infrastructure integrated thousands of computers and storage systems in hundreds of data centers. The infrastructure provided global computing resource for storing and processing the petabytes of data generated by the LHC. Based on the middleware packages developed for the EDG project, a new middleware software package known as LCG was developed. Thus, the foundations of the European Grid Infrastructure were placed.

As a continuation of the work of the two projects in 2004, Enabling Grids for E-Science in Europe (EGEE) project [5] was launched. Undoubtedly, this project had the largest contribution to the development of the European Grid Infrastructure with respect to the provided hardware and software capabilities. A new distribution based on the software packages of EDG and LCG environments was created. The name of the new middleware was gLite.

After the end of EGEE project in 2010, the middleware components in gLite became part of the EMI distribution. The European Middleware Initiative (EMI) [6] collaborate the major European middleware providers as ARC, dCache, gLite and UNICORE, in order to deliver a consolidated set of middleware components for deployment in grid infrastructures. EMI is the current middleware of the EGI.

In 2013 the EMI project ends and currently there is not new project specially dedicated for grid middleware development. This was defined [7] as a milestone for EGI and grid computing in Europe and brings the evolution of the EGI infrastructure to the new turning point. For EGI, the answer is the EGI Platform Architecture – a new platform-based architecture which to allow any type and any number of community platform to co-exist on the same physical infrastructure. This new architecture allows preservation of the current EGI Core Infrastructure and establishment of new EGI Cloud Infrastructure.

For the Grid middleware however, there is not outlined future trends of development. During the different projects the middleware software packages evolved through: EDG middleware, LCG middleware, gLite middleware to EMI middleware. For the last two grid environments – gLite and EMI, some service-oriented approaches were considered. Unfortunately, the idea for the service-orientation was partially accomplished.

2. EGI and the Service-Oriented Architecture

Service-Oriented Architecture (SOA) [8] is an architecture style targeted to create loosely coupled interactions among services. Services provide object-oriented capsulated business logic or functions of IT components. The interactions between services are usually implemented by Web services. The standards for Web services and Service-Oriented Architecture are developed by OASIS and W3C.

The standards in Grid are developed by Globul Grid Forum (GGF). GGF intends to define Grid specifications that can become broadly accepted standards for the international society to exchange ideas, experience and best practices. Open Grid Services Architecture (OGSA) is the GGF solution for information and resource sharing among organizations, which utilize products from different vendors.

The development of the EGI's Grid middleware in the different projects has followed different software approaches. In the very beginning there was an attempt to apply the standards introduced by GGF. Subsequently in the next Grid middleware the direction has been directed to the Service-oriented architecture. The current Grid middleware distance itself from the idea of service-orientation.

For the Grid environments gLite and EMI, some service-oriented approaches were considered. The architecture of gLite and its services are presented in [9] as services which follow the Service-Oriented Architecture. This is pointed by the author as feature which "will facilitate interoperability among Grid services and allow easier compliance with upcoming standards, such as OGSA". All these ideas were laid down into the environment's architecture. Unfortunately they were partially accomplished.

Some analyses in this direction were made by the author in [10]. The lack of discovery services and services for composition into gLite middleware, made the

environment partially service-oriented. Even more, there is lack of widely accepted mechanisms for business process orchestration, mediation and monitoring. In [11] and [12] the author presents a solution in the context of SOA for business process orchestration, mediation and monitoring for the European Grid Infrastructure with gLite middleware.

Despite of all the effort of the developers to make the g-Lite middleware service-oriented, some legacy components of the environment remained unchanged. This influenced the architecture of the next middleware successor of gLite – EMI.

Learn from the experience of gLite, EMI architecture was defined in [13] as "a set of features and services released together that tackles a set of requirements and can handle a set of use cases", where the term service is used in a broader sense than web service. This definition clearly shows the alienation of the EMI middleware from the Service-oriented Architecture. With the end of the EMI project and without new dedicated project for grid middleware development, the EMI middleware will stay in the current state as partially service-oriented environment.

With the end of the EMI project, the EGI community lost one of its two largest technology providers and contributors to the Unified Middleware Distribution. This brings the evolution of the EGI infrastructure to the new turning point. For EGI, the answer is the EGI Platform Architecture - new software integration and provisioning process, based on cloud computing.

3. The European Grid Infrastructure from Grid to Cloud

Gartner [14] defines cloud computing as "a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using Internet technologies". Cloud capabilities are defined as services. Such are Computing as a Service (CaaS), Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS).

The main Grid idea is Computing as a Service (CaaS), which means that the user has not to care where his data resides and which computer executes his request. He will be able to request information and received it timely in accordance with his requirements.

A Platform as a Service (PaaS) provides user an application development environment delivered over a network, typically the Internet. It is a broad collection of application infrastructure (middleware) services.

The Infrastructure as a Service (IaaS) provides user an infrastructure of compute resources, complemented by storage and networking capabilities on-demand. The user is able to self-provision this infrastructure, using a Web-based graphical user interface for the overall environment.

A Software as a Service (SaaS) provides the user standardized, network-delivered IT applications. The service provider retains ownership of the software and hardware assets and handles all maintenance and support. The user is able to add and subtract licenses on-demand. The mechanism of representing applications as services enables module reuse. It also enables dynamic binding of services. All these are features of the service-oriented architecture.

The service-oriented architecture offers an agile application infrastructure that can meet the requirements of cloud. Therefore, the service-oriented approach is the preferred approach for building cloud systems. Cloud computing provides standardized and easy access to shared hardware and software at low cost. Together SOA and cloud can provide a complete services-based solution. Example application of SOA in cloud is shown in [15], where the authors proposed service-oriented generic resource framework for cloud systems,

which represents datacenter resources in a uniform way, allowing generic administration without knowledge of the underlying resource access protocol.

To provide generic, consistent and flexible access to EGI resources, EGI establish new software integration and provisioning process, based on cloud computing or more precise on a locally deployed IaaS Clouds (Figure 1).

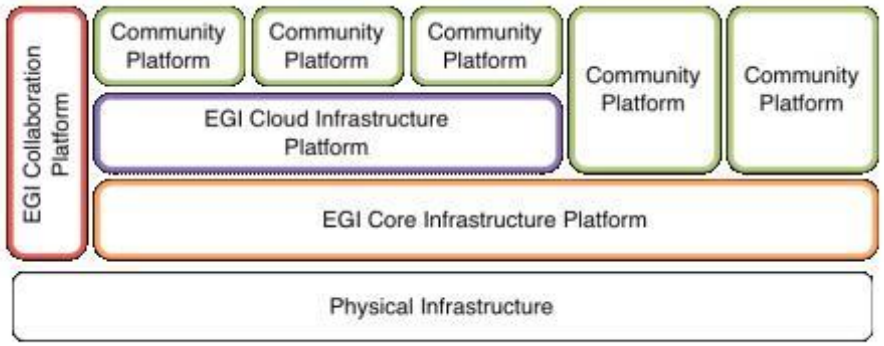


Figure 4. The EGI Platform Architecture [16].

The Core Infrastructure Platform [16] is organized in two distinct platforms: The EGI Core Infrastructure Platform and the EGI Cloud Infrastructure Platform. The EGI Core Infrastructure Platform includes only services that are necessary for operational federated distributed computing infrastructure.

The EGI Cloud Infrastructure Platform is deployed on the top of the EGI Core Infrastructure Platform. The new cloud infrastructure will provide new federation and distribution services directed to the cloud computing and will supports new research communities who wish to deploy their own Virtual Research Environment.

The EGI Collaboration Platform will offer services that allow collaboration across research communities and their domain-specific community platforms. It will encapsulate services that are common across multiple communities and are not critical to the operation of the EGI production infrastructure.

4. Conclusions

It becomes clear that the evolution of EGI infrastructure is directed to the cloud computing. However, what software approaches will be used for the realization of this idea at this point is not so clear. Because of the inhomogeneity of the EMI middleware and the evolution to the cloud the service-oriented approach seems to be the right direction. The lack of new dedicated project for grid middleware development however can influence the choice of approaches for realization.

5. Acknowledgements

Research in this paper is funded by Bulgarian Science Fund under contract DFNI-I01/12 "Contemporary programming languages, environments and technologies and their application in building up software developers".

References

1. Kesselman C. and I. Foster, The Grid: Blueprints for a new Computing Infrastructure, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999, ISBN: 1-55860-475-8
2. European Grid Infrastructure (EGI), http://www.egi.eu
3. European DataGrid Project (EDG), http://eu-datagrid.web.cern.ch/eu-datagrid
4. Worldwide LHC Computing Grid (WLCG), http://lcg-archive.web.cern.ch/lcg-archive/public
5. Enabling Grids for E-Science in Europe (EGEE), http://egee-technical.web.cern.ch/egee-technical
6. European Middleware Initiative (EMI), http://www.eu-emi.eu/
7. EMI Issues, http://www.egi.eu/news-and-media/newsletters/Inspired_Issue_11/ platform_architecture.html
8. Erl T., Service-Oriented Architecture: Concepts, Technology, and Design, Prentice Hall Publishers, 2005, ISBN: 978-0131858589
9. Laure, E., et al., Programming the Grid with gLite, Computational Methods in Science and Technology 12(1), 2006, pp. 33-45, http://www.cse.hcmut.edu.vn/~ptvu/gc/gLite.pdf
10. Goranova R. D., Service composition tools in g-Lite, Conference Proceedings of the Fifth International Conference ISGT, 2011, pp. 228-235, ISBN 978-0-7354-0976-7, ISSN 0094-243X.
11. Goranova, R. D., Framework for service composition in g-Lite, American Institute of Physics, Conference Proceedings Volume 1404, 2011, pp. 218-224, ISBN 978-0-7354-0976-7, ISSN 0094-243X.

Formalization of WS-BPEL Business Process in CSP

Vladimir Dimitrov

Faculty of Mathematics and Informatics, University of Sofia,
5 James Bourchier Blvd., 1164 Sofia, Bulgaria
cht@fmi.uni-sofia.bg

Absract. WS-BPEL is widely accepted standard for specification of business processes. It is based on formal models of process algebra and Petri nets. As result of that it is easy to write business processes with deadlocks and other unwanted features. This paper shows how an example WS-BPEL process can be converted to formal CSP model. The last one can be formally verified.

Keywords. Business process, communicating sequential processes, formalization, verification.

1. Introduction

WS-BPEL [1] supports two kinds of business processes: executable and abstract ones. Behavioral semantics of executable business processes is well defined in the notation framework. The only problem is with the extensions, because they go outside the notation framework.

WS-BPEL abstract business processes intention is to describe Web services interactions without internal implementation details. The standard defines Basic Executable Completion for abstract processes and requires for every abstract process to exist at least one executable process that confirms this rules. The problem is that it is possible to be generated an executable process with the structure of abstract one but with different behavior. The abstract process represents a class of executable processes compatible with it. As result of that it is more productive to verify abstract business processes. The last ones do not implementation details that have no impact on Web services interactions.

WS-BPEL process could be viewed as Web service implemented in WS-BPEL code. It is not meaningful to formalize WSDL [2] specifications of Web services, participating in the interactions on the business process (including the business process as a Web service), because they are interfaces without any behavior. It is significant to verify WS-BPEL code of the business process.

The idea of this paper is to verify WS-BPEL code converting it into CSP [3] model. An example business process, taken from the standard, is used to illustrate this modelling procedure.

2. Formalization of the WS-BPEL Specification

For specification and verification purposes is used PAT [4] and its version of CSP.

The business process logic written is pseudo code is:

```
receive shipOrder
if condition shipComplete
  send shipNotice
else
  itemsShipped := 0
  while itemsShipped < itemsTotal
    itemsCount := opaque
    // non-deterministic assignment corresponding e.g. to
    // internal interaction with back-end system
    send shipNotice
    itemsShipped = itemsShipped + itemsCount
```

The process is instantiated when a shipOrder is received. If this order has been executed yet then a shipNotice is replied. This information is contained in the message header property shipComplete. Otherwise, a cycle is executed for order execution. At every step, part of the items are delivered and a notification is send. The counter is incremented with the number of sent items. The cycle exits when all items are delivered and then the process is terminated. In the abstract process, the number of delivered items at every step is non-deterministic. This information is retrieved from the back end system. From interactions point of view, this process is very simple. Initially, the process waits to receive an order message from a consumer and then replies with one or more messages. There are no fault handlers, no compensators, no return values.

The example abstract business process starts with a heading:

```
<process name="shippingService"
  targetNamespace="http://example.com/shipping/"
  xmlns="http://docs.oasis-open.org/wsbpel/2.0/process/abstract"
  xmlns:plt="http://example.com/shipping/partnerLinkTypes/"
  xmlns:props="http://example.com/shipping/properties/"
  xmlns:ship="http://example.com/shipping/ship.xsd"
  xmlns:sif="http://example.com/shipping/interfaces/"
  abstractProcessProfile=
    "http://docs.oasis-open.org/wsbpel/2.0/process/abstract/ap11/2006/08">
  <import importType="http://schemas.xmlsoap.org/wsdl/"
    location="shippingLT.wsdl"
    namespace="http://example.com/shipping/partnerLinkTypes/" />
  <import importType="http://schemas.xmlsoap.org/wsdl/"
    location="shippingPT.wsdl"
    namespace="http://example.com/shipping/interfaces/" />
  <import importType="http://schemas.xmlsoap.org/wsdl/"
    location="shippingProperties.wsdl"
    namespace="http://example.com/shipping/properties/" />
```

This heading is simply declaration and does not contain any behavioral information. It imports needed WSDL specifications of interacting Web services.

The business process body begins with a partner link definition:

```
  <partnerLinks>
```

```
  <partnerLink name="customer"
    partnerLinkType="plt:shippingLT"
    partnerRole="shippingServiceCustomer"
    myRole="shippingService" />
 </partnerLinks>
```

The role of the business process is a shippingService. The link could be modelled in CSP as a channel:

channel customer 0;

This channel has no buffers, i.e. only one message can be exchanged through it. The channel in PAT could have some capacity, but in classical CSP it is modelled. There is no need of channel buffering.

Three variables are defined:

```
 <variables>
  <variable name="shipRequest" messageType="sif:shippingRequestMsg" />
  <variable name="shipNotice" messageType="sif:shippingNoticeMsg" />
  <variable name="itemsShipped" type="ship:itemCountType" />
 </variables>
```

The first two variables are buffers for receiving and sending messages. The third one is a cycle counter. The process state is represented by them. This business process has the states: initial (waiting for an order), check order, and execute order. These states could be modelled as three subprocesses as shown below.

Then a correlation set is defined:

```
 <correlationSets>
  <correlationSet name="shipOrder" properties="props:shipOrderID" />
 </correlationSets>
```

The shipping order ID is used for dialog support. This process can be transformed to stateless process if we accept that all the process state information is contained in one document, for example shipOrder message. Then this message would contain all needed information, including order ID. The example process does not receive more than one message – it only sends messages to the consumer. Only, the consumer has to distinguish received messages from the shipping service. In this case, there is no need to coordinate process instances with the correlation set: when a main process instance is started, it simply deliver shipOrder to its subprocesses instances as a parameter.

Then the process activity is a sequence. The first it activity is to receive a shipping order from the partner link:

```
  <receive partnerLink="customer" operation="shippingRequest" variable="shipRequest">
   <correlations>
    <correlation set="shipOrder" initiate="yes" />
   </correlations>
  </receive>
```

This is modelled in the main process as a channel read:

customer?shipOrder

The main process would start subprocesses for to check and possibly to execute the order, but at the same time it would restart recursively parallel for new order:

shippingService() = customer?shipOrder -> (checkOrder(shipOrder) ||| shippingService());

WS-BPEL code then checks the order in an “if” with condition:

```
   <condition>
    bpel:getVariableProperty('shipRequest', 'props:shipComplete')
```

```
    </condition>
```

Outputs of this check could be modelled with two events: shipComplete and shipNotComplete.

The "then" part of the "if" activity simply notifies the consumer:

```
    <sequence>
      <assign>
        <copy>
          <from variable="shipRequest" property="props:shipOrderID" />
          <to variable="shipNotice" property="props:shipOrderID" />
        </copy>
        <copy>
          <from variable="shipRequest" property="props:itemsCount" />
          <to variable="shipNotice" property="props:itemsCount" />
        </copy>
      </assign>
      <invoke partnerLink="customer"
       operation="shippingNotice"
       inputVariable="shipNotice">
       <correlations>
         <correlation set="shipOrder" pattern="request" />
       </correlations>
      </invoke>
    </sequence>
```

This is modeled in the subprocess:

```
checkOrder(shipOrder) =
 (shipComplete -> customer!shipOrder -> Skip) []
 (shipNotComplete -> executeOrder(shipOrder));
```

The "else" part – execute the order in a cycle:

```
    <sequence>
      <assign>
        <copy>
          <from>0</from>
          <to>$itemsShipped</to>
        </copy>
      </assign>
      <while>
        <condition>
          $itemsShipped &lt; bpel:getVariableProperty('shipRequest', 'props:itemsTotal')
        </condition>
        <sequence>
          <assign>
            <copy>
               <opaqueFrom/>
              <to variable="shipNotice" property="props:shipOrderID" />
            </copy>
            <copy>
              <opaqueFrom/>
              <to variable="shipNotice" property="props:itemsCount" />
```

```
          </copy>
        </assign>
        <invoke partnerLink="customer"
         operation="shippingNotice"
         inputVariable="shipNotice">
         <correlations>
          <correlation set="shipOrder" pattern="request" />
         </correlations>
        </invoke>
         <assign>
          <copy>
           <from>
            $itemsShipped + bpel:getVariableProperty('shipNotice', 'props:itemsCount')
           </from>
           <to>$itemsShipped</to>
          </copy>
         </assign>
        </sequence>
       </while>
      </sequence>
```

The "while" cycle is modelled as subprocess:

```
executeOrder(shipOrder) =
 (itemsShipped -> Skip) []
 (itemsNotShipped -> change_itemsCount ->
  customer!shipOrder -> executeOrder(shipOrder));
```

The subprocess exits when all items are shipped. Otherwise it notifies the consumer and recursively executes itself. Here, checks again are modeled with events.

The whole CSP model is:

```
channel customer 0;
shippingService() =
 customer?shipOrder -> (checkOrder(shipOrder) |||
 shippingService());
checkOrder(shipOrder) =
 (shipComplete -> customer!shipOrder -> Skip) []
 (shipNotComplete -> executeOrder(shipOrder));
executeOrder(shipOrder) =
 (itemsShipped -> Skip) []
 (itemsNotShipped -> change_itemsCount ->
  customer!shipOrder -> executeOrder(shipOrder));
var count = 10;
shippingServiceCustomer() =
 if (count > 0) {customer!count -> {count--} -> receive()}
  else {Skip};
receive() = customer?shipNotice -> receive();
System() =
 shippingServiceCustomer() ||| shippingService();
#assert System() deadlockfree;
```

For verification purposes two additional subprocesses are added: shippingServiceCustomer and System. An PAT verification assert is added at the end.

3. Conclusion

WS-BPEL specification of business process specifies its behavior. Its formalization in CSP is maximally abstracted many implementation details saving the original interaction flow. The CSP specification can then be formally verified. The CSP model is very compact and readable.

This example of formalization demonstrates an approach to formal verification of business processes. Full description of this research approach will be published in future.

4 Acknowledgements

Research in this paper is funded by Bulgarian Science Fund under contract DFNI-I01/12 “Contemporary programming languages, environments and technologies and their application in building up software developers”.

References

1. OASIS, Web Services Business Process Execution Language Version 2.0, OASIS Standard, 11 April 2007, http://docs.oasis-open.org/wsbpel/2.0/wsbpel-v2.0.html
2. W3C, Web Services Description Language (WSDL) 1.1, W3C Note, 15 March 2001, http://www.w3.org/TR/wsdl
3. C. A. R. Hoare, Communicating Sequential Processes, Prentice Hall International. ISBN 0-13-153271-5, 1985.
4. L. Yang, PAT: Process Analysis Toolkit: An Enhanced Simulator, Model Checker and Refinement Checker for Concurrent and Real-time Systems, http://www.comp.nus.edu.sg/~pat/.

The Architecture of Agent-Oriented Situation Management System

Oleksii E. Kovalenko

IMMSP - NASU, Acad. Glushkov Avenue, 42, 03680 Kyiv, Ukraine
koval@immsp.kiev.ua

Abstract. The paper reviews architecture principles of situation management systems design based on agent-oriented approach. The architecture for agent-oriented situation centers based on evolutionary functional principle is proposed. The elements and specifics are defined for proposed model. Common mechanism of formal knowledge application for multiagent system management is defined. Promising directions for future research are outlined.

Keywords. Situation management system, decision support system, agent-oriented system, ontology, knowledge model, infocenose.

1. Introduction

Decision making in complex systems is multifactor multivariant task and usually occurs with the use of information systems of a special class - decision support systems (DSS). Many of the tasks of public administration, forecasting, analyzing and emergencies recovering are interdisciplinary and intersectoral nature are solved using the collaborative work of groups of experts and analysts on the technological basis of DSS of Situation Centre type ("Situation Room") [1, 2].

Situation Centers (SC) are intended for management decisions concerning complex multi-object structure of administration in which the poorly structured and poorly formalized, input data about objects of ambiguous, incomplete and contradictory and cost of the errors in the solution is very high. [1]. Thus the definition of architectural, functional, organizational principles of the creation and development based on these principles, components of SC is an actual task.

2. The analysis of the current state of research and publications

Features and general principles of the SC described in [1, 2, 3]. How organizational control, situational control is characterized by such features [3]:

- Uniqueness;
- Lack of formalized objective of existence;
- No classical definition of the optimality;
- The dynamic nature;
- Incomplete description;
- The presence of agency.

In the development of SC used four groups of principles [1]:

- Architectural and functional;
- Organizational and technological;
- Information;

- Intelligence.

In each of these groups, the principles put forward their demands to the SC components that ensure its operating characteristics. Diversity and variety of requirements and the need to adapt operation procedures to the specific tasks of SC using determines the use of flexible solutions for SC design, particularly with the use of object-oriented [4], functional [5] and multi-agent approaches [6, 7].

3. Statement of the problem

Research purpose is the development and justification of knowledge managed multi-agent model of collective decision making under SC. SC requires a multidiscipline management procedures to ensure the adaptation of its operations for solving specific problems of decision support. Such adaptations may be made by applying the models of the domain knowledge for management of specific decision making.

4. The main problem

Decision making in SC is working out the workflow, based on modern information technology to ensure the use of formal and informal knowledge bases for selecting control actions on the object, a process subject area during the period of time in order to achieve the object, the process, the problem area necessary condition.

General outline of decision making schema in SC is depicted on Fig. 1.

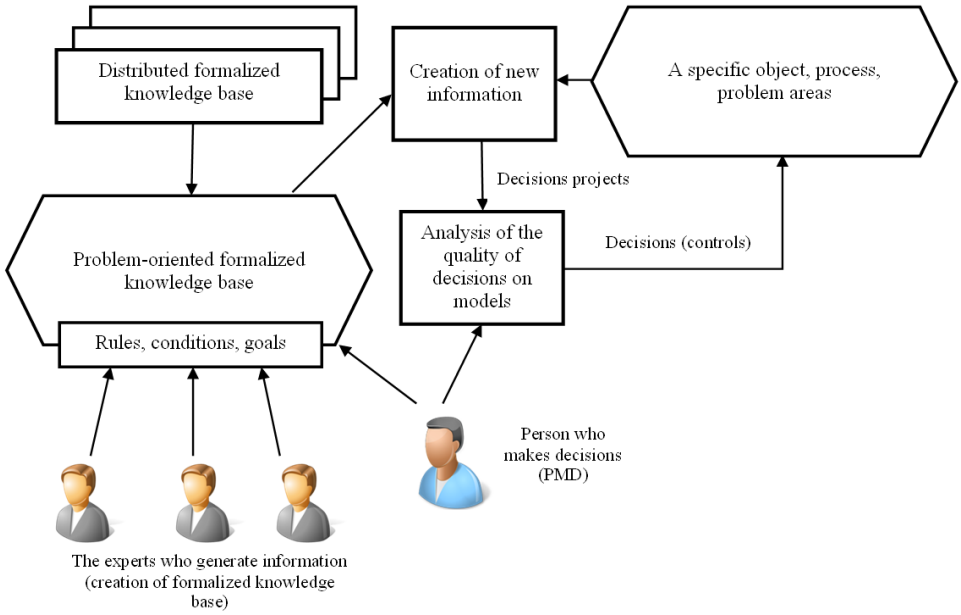


Fig. 41 – Decision making schema in SC

Thus it is clear that SC should be equipped with a special set of software and hardware for personal and collective analysis of the group heads.

One of the main features of the SC functioning is the decision making in the mode of teams collective work, which includes a person who makes decisions (PMD), a group of experts, a group of analysts and staff of SC [1]. Methodological foundations of multipurpose SC, as a purposeful organizational structures, involves identifying those aspects of [8]:

1) mission statement - generalized objective, formulated in the context of a strategy of SC;

2) vision statement - assessment of environment and a clear definition of criteria and perspectives;

3) management strategies (strategic management) as a continuous process associated with the formulation of organizational solutions for an efficient and productive activities in the context of the mission, vision and global business objectives;

4) tree objectives (goals, objectives, targets) - hierarchical set of final states or outcomes to attain the aims of SC activity;

5) policies - the institutional arrangements to achieve of activity objectives;

6) patterns of activity (activity models) - a set of methods, tools and algorithms that define and describe the operating procedures at all levels of organizational structure.

Designated aspects of the corresponding functional characteristics of SC - a set of services (products and services). The scope and quality of information services is subject to available information resources - databases, knowledge bases, repositories of standard procedures and precedent decisions. Thus, for constructing collective decision-making model it is necessary to make a decomposition of the models in accordance with its principles and aspects.

The variety of tasks to be solved in SC must meet the technological level of SC maturity for their solving. Therefore it is necessary to automate the preparing and adapting SC to specific tasks of decision making taking into account all aspects and activities. This problem can be solved by using an evolutionary approach using problem-oriented knowledge bases that describe the information and functional environment of collective decision-making.

So the functional environment of collective decision-making must be supported by appropriate architectural solutions.

In appliance with decision making schema depicted on fig. 1 architectural components of SC include [9]:

- Interface subsystem;
- Subsystem for recognition and estimation of the state of situations;
- Subsystem for retrospective analysis of decisions taken and learning;
- Subsystem for production of alternatives based on knowledge;
- Subsystem for production of alternatives based on macro models;
- Subsystem for collective estimation, decision making, goal setting subordinates.

These subsystems are, in turn, includes:

- Interface means;
- Knowledge base about problem area and its environment;
- Knowledge base about experts in specific areas;
- Knowledge about rules, conditions and rules of decision making process;
- Means and engines of inference on knowledge;
- Means for simulation and visualization;
- Means of supporting collective decision-making process;
- Tools for documenting and archiving decision-making process;
- Security Tools etc.

Modern information technologies widely use service oriented approach, virtualization and software agents based intellectualization. So, the SC architecture may be represented as a set of virtual components which integration is based on the interaction of agents. Means and

tools of decision making process during situation management should be accessible as services.

5. Elements of agent-oriented model of decision-making in SC

An intelligent agent (IA) is an autonomous entity which observes through sensors and acts upon an environment using actuators (i.e. it is an agent) and directs its activity towards achieving goals (i.e. it is rational) [10]. So, generalized model of agent is a tuple $A = <S, D, B>$, where S is a set of sensors, D is a set of actuators, B is a behavior model of agent. Agent's behavior is described by the agent function f_A that maps any possible percepts sequence P^* to a possible action B_A [10, p. 35] or mathematically speaking:

$$f_A : P^* \rightarrow B_A .$$

Intelligent agents may also learn or use *knowledge* to achieve their goals. A variety of types of agents determines the multiplicity of their representations models. One of the most advanced software agent's models is the BDI-model (belief–desire–intention model).

Multi-agent system (MAS) - a system where multiple agents can the exchange of current information (interact). Multi-agent system is described by a model [7]:

***MAS* = (*A, E, R, ORG, ACT, COM, EV*)**,

where ***A*** - the set of agents that can operate in certain environments ***E***, which are in certain respects ***R*** and which interact with one another, forming a particular organization ***ORG***, those that have a set of individual and joint action ***ACT*** (behavior strategies and deeds), including the possible communication actions ***COM***, and is characterized by capabilities of evolution ***EV***.

The very agent in ***MAS*** is characterized by the following properties:

- knowledge - permanent part of the agent's knowledge about themselves, about the environment and about other agents, this knowledge does not change during its operation;
- beliefs - agent knowledge about the environment, particularly about other agents, this information can change in time, become even false, but the agent might do not to know about it and continue to believe that conclusions may be based on such knowledge;
- commitments with respect to other agents - tasks that the agent undertakes to request (order) of other agents within cooperative goals.
- desires - states, the achievement of which corresponds to the appointment of an agent;
- intentions - what agent has done, and what follows from its obligations or aspirations;
- objectives (goals, objectives, targets, aims) - particular set of conditions, achievement of which is the meaning of the current strategy of agent behavior;

and also:

- mobility - agent's ability to migrate, to move the network to find the necessary information to solve their problems in cooperative solving them together or with other agents;
- benevolence - the readiness of agents to help each other and readiness to solve exactly the task that has commissioned by the user;
- veracity - agent property does not manipulate false information, knowing that it is false;

– rationality - ability to act such as to achieve their goals, rather than avoid it, at least within its knowledges and beliefs.

Agents in MAS must interact in accordance with the decision making process used in the SC. Models of decision-making procedures are stored in the knowledge base (KB) SC, which is replenished and updated during the operation of the SC with the experience of previous decision-making. Thus the functioning of SC has an evolutionary character. KB SC is a hierarchy of domains of knowledge about different aspects of the functioning of SC procedures, which include the knowledge relating to specific areas of situational management, and knowledge related to the operation regulations of SC, including knowledge of the organization process of collective decision making. This knowledge is used to determine the list of required agents and patterns of agents behavior and interaction during the operation of SC.

Considering hierarchy of knowledge in the SC agents can be classified by their purpose and use. To date, proposed a large number of models of intelligent agents, each of which has its advantages and disadvantages [7]. Therefore, the choice of a specific set of agents for the implementation of certain procedures in the SC should be carried out by selection of agents in this set. For this purpose knowledges about the agents and their properties from the KB of available agents are used.

Considering algorithm for collective decision-making the implementation of each stage of this process is supported by relevant groups ("colonies") agents so that the results of the colony activity are inputs for other colony activity, and eventually obtained the final outcome of collective decision-making process in a specific format. Thus, the decision is the result of superposition of functioning of the agents colonies set on the given information field. Thus, there is a problem of formal definition of specifications for decision making problems that is determined the initial composition of the "agents world" for a particular decision making problem. Based on the specifications is determined the composition of each colony, selected from the "population" of much alike agents among themselves.

Thus, the creating an environment for collective decision-making in SC is a task of infocenosis creating (information "biocenosis") [11] as a special type agents technocenosis [12] and associated with the formation of separate populations to ensure the objective function of the system. With the initial conditions of uncertainty and decision-making tasks specifications followed uncertainty populations and agents biocenosis. Their specification can be made by analyzing the problem specification and analysis of available "world" agents. Building the "agents world" on taxonomic characteristics and maintaining information on taxonomy of ontological KB provides the possibility of formal association of agents to resolve problems making decisions based on given specifications and constraints. As agents of change in the evolution of the system, these changes can be described by agents “mems” (informational "genes" - steady information structures, capable for replication). Then, continuing the biological analogy, we can assume that the effective functioning of agents infocenosis depends on the level and content of symbiotic relationships between “mems” inside the agents and agents themselves.

Taking into account the preceding reasons for creating of infocenosis for adaptive agent-oriented SC is proposed the tuple of agents populations:

$$\boldsymbol{M} = <\boldsymbol{M}_i \mid i = 1\ldots n>,$$

where i – index (name) of population; n – infocenosis power.

For example, for MAS of collective decision-making is offered the following populations tuple:

$$\boldsymbol{M} = <\boldsymbol{M_{Env}}, \boldsymbol{M_{eXpert}}, \boldsymbol{M_{DecProc}}, \boldsymbol{M_{Rest}}, \boldsymbol{M_{Crit}}, \boldsymbol{M_{Dec}}, \boldsymbol{M_{Goal/Obj}}, \boldsymbol{M_{Brunch}}>,$$

where $\boldsymbol{M_{Env}}$ – population of environment agents; $\boldsymbol{M_{eXpert}}$ - population of expert agents; $\boldsymbol{M_{DecProc}}$ - population of decision making procedures; $\boldsymbol{M_{Rest}}$ - population of agents-restrictors; $\boldsymbol{M_{Crit}}$ - population of criteria; $\boldsymbol{M_{Dec}}$ - population of decision precedents agents; $\boldsymbol{M_{Goal/Obj}}$ - population of objective designations agents; $\boldsymbol{M_{Brunch}}$ - population of models agents for decision making domain.

Agents possessing of more accurate knowledges about the environment and have adequate behavior algorithms as well as in nature, are more likely to preserve the environment and obtaining resources to fulfill its functions. Obtaining the new knowledge and modification together with the other factors increase the chances of preserving agents. To transfer knowledge and modifying agents used steady information structures - *mems*. Thus MAS characterized by a certain "social" behavior.

To save *memplex* (mems group, characterized by a symbiotic relationship, a set of ideas that reinforce each other) that characterize the behavior of agents can be used ontological model. Stable memplex combination forms typical construction of knowledge - knowledge patterns. These knowledge patterns can be described with the use of ontolology model. Knowledge pattern should include such models as:

- Roles model;
- Communication model;
- Services model;
- Contacts model;
- Rules model;
- Restriction model;
- Criteria's model.

The above models also may be represented as second level ontologies.

Simple formal model of ontology is the $\boldsymbol{O}=<\boldsymbol{T}, \boldsymbol{R}, \boldsymbol{F}>$, where $\boldsymbol{T}$ is the set of the terms of problem area, $\boldsymbol{R}$ is the set of relations between the terms from $\boldsymbol{T}$, $\boldsymbol{F}$ is the interpretation functions defined on terms from $\boldsymbol{T}$ and/or relations from $\boldsymbol{R}$. So, the multilevel ontology is

$$O = \left\langle O_i^j, R^{j-1}, R^{j+1}, F^{j-1}, F^{j+1} \right\rangle,$$

where O_i^j fragment i from ontology of level j, R^{j-1} and R^{j+1} relations between current level of ontology and upper/lower layer, F^{j-1} and F^{j+1} interpretation functions defined for terms and/or relations upper/lower levels. Virtual architecture of agent oriented situation management system will be constructed as a set of temporal actual relations in the control ontology of MAS. The knowledge of particular agent's population will be determined as appropriate fragment of problem area knowledge $O_i = M_i(O)$.

Further researches are focused on building metamodel patterns of knowledge base for SC and optimizing of its semantic volume and computational complexity.

6. Conclusion

Specific of SC requires the creation of adequate models describing their architecture. As this model uses MAS model. Properties of agents and their interaction in MAS with signs of "social behavior". Taking into account it is natural to consider MAS for collective decision-making in the SC as infocenosis, consisting of populations of agents for different purposes, and are characterized by information-functional symbiotic relationship. Proposed

as a criterion of evolutionary selection of agents to use the level of adequacy of their knowledge about the environment and subject fields of decision-making and consumption of resources. As a mechanism for the preservation of knowledge proposed to use the ontological model of knowledge based on steady semantic construction - mems and memplex that are formed and improved in the process of evolution of MAS for SC.

References

1. Morozov A.O. Situation centres - the base of control by the great dimension organizational systems // Mathematical Machines and Systems, 1997, 2, (7-10).
2. Raikov A.N. Situation room for corporate decision support // Open Systems Publications, 1999, 7-8. P.56-66. (http://www.osp.ru/os/1999/07-08/179889/)
3. Pospelov D. A. Situational control: Theory and practice, Nauka Publishers, Moscow, 1986.
4. Booch Grady, Maksimchuk Robert A., Engel Michael W., Young Bobbi J., Conallen Jim, Houston Kelli A. *Object-Oriented Analysis and Design with Applications*, Addison-Wesley Professional, 2007.
5. Kovalenko Oleksii E. Functional Approach in the Simulation Modeling of the Problem-Oriented Information Systems // Комуникации, информационни технологии и статистика. Актуални проблеми на теорията и практиката. - София : Университетско издателство "Стопанство", 2011. - С. 278-281.
6. Shoham Yoav, Leyton-Brown Kevin, Multiagent systems: algorithmic, game-theoretic, and logical foundations, Cambridge University Press, New York, 2009.
7. Wooldridge M. J. An Introduction to Multiagent Systems, John Wiley and Sons, Ltd., Chichester, 2002.
8. Hills Ch., Jones G. Strategic Management, Houghton Mifflin Company, New York, 2008.
9. Kusmenko G.E., Pleesh V.E. Functional architecture of the integrated system of the decision making support in the situation centre conditions // Mathematical Machines and Systems, 1997, 1. (56-63).
10. Russell, Stuart J.; Norvig, Peter. *Artificial Intelligence: A Modern Approach* (3rd ed.), Upper Saddle River, New Jersey: Prentice Hall, 2010.
11. Leonov V.P. Camouflage Memes of Scientific Schools Infocenosis (http://www.biometrica.tomsk.ru/phil_math.htm)
12. Gnatiuk V. I. The Law of Technocenosis Optimum Construction (http://www.gnatukvi.ru/index.files/zakon.pdf)

IMPACT OF BUSINESS PROBLEM CHARACTERISTICS ON THE ARCHITECTURE AND SPECIFICATION OF INTEGRATION FRAMEWORK

S. Kouzmanov[(1)] , Monika Tsaneva[(2)]

[1] Department of ITC, UNWE, 1700 Sofia, Bulgaria
skouzmanov@gmail.com

[2] Department of ITC, UNWE, 1700 Sofia, Bulgaria
mtzaneva@unwe.bg

Absract. Modern enterprise information environments usually host a lot of different software systems that are involved in a wide range of business tasks and operations. For a successful usage of the whole heterogeneous information environment of an enterprise, the integration framework used to connect different software systems should take into account the specifics of the business problem and scope. Different needs can vary a lot, but still they can be generalized into larger groups of common integration and data processing problems. This can be used as a foundation for researching the impact of business problems and needs on an integration frameworks used into and enterprise. This paper uses a brief, general classification of business problems and proposes a mapping between business needs and integration approaches.

Keywords. Integration, real time processing, batch processing, services, data integration

1 Introduction

Nowadays the implementation and execution of an enterprise business is becoming more and more complex. With the process of globalisation, expansion to new markets, introducing new products and services and also with the advance of business oriented information technology, organisations increased their need for fast and accurate information processing. This includes an increase in data volumes to be processed and also a diversification in types of business problems and needs.

On the other hand, modern organisations have complex, heterogeneous information environments and application infrastructures. This leads to a constant need for development of an enterprise-wide integration frameworks [5]. Since various applications that should be interconnected by the integration framework are deployed in different business problem context, the way they are integrated with other applications can be affected by the type of business needs they solve.

2 Different Business Problems and Needs

Recipients of information within an organisation can be different and their needs can vary. This changes the way, the time, the format and the representation of similar core data delivered to different addressees.

For example, online customers will need accurate information about an item they intend to purchase – is the item in stock, how many units of it is remaining, what's the exact current price. Operating staff, like employees who are in charge for order shipping will also need real time data about customers who are waiting their orders to be shipped. Managers or marketing staff will also need real-time data about sales, quantities and so on in order to adjust promotions, marketing initiatives and so on. All of the above examples can be generalized as core business and customer support data. If this type of data is processed in a real time manner the organisation will be more flexible, it will deliver more value to customers and new business opportunities will be enabled.

On the other side, there is also data that is not needed in a real time or near real time fashion. For example a top management member or a board of directors will not need information about today orders, but will be interested in the trend of sales for the last quarter or in a prediction for the sales of a product for the next year. Such information can be provided after collecting and analysing of large volumes of data. Such big data can be transferred to a business intelligence system in the so called batch window, when the core business is not fully loaded and has free resource for synchronising with the BI. Another example is some bank transactions which are not due on the same day. Data about such kind of transactions can be transferred between systems during the night, when information infrastructure is not loaded.

According to the above examples we can suppose that in the aspect of their data integration, most of the various business needs can generally be classified and beaddressed as either as core business, real-time data integration or as non-real time, batch data replication problems.

3 Real Time and Batch Data Processing

After we identified the possible classification of different business problems we can examine the difference between batch and real time data processing as a fundament of further research.

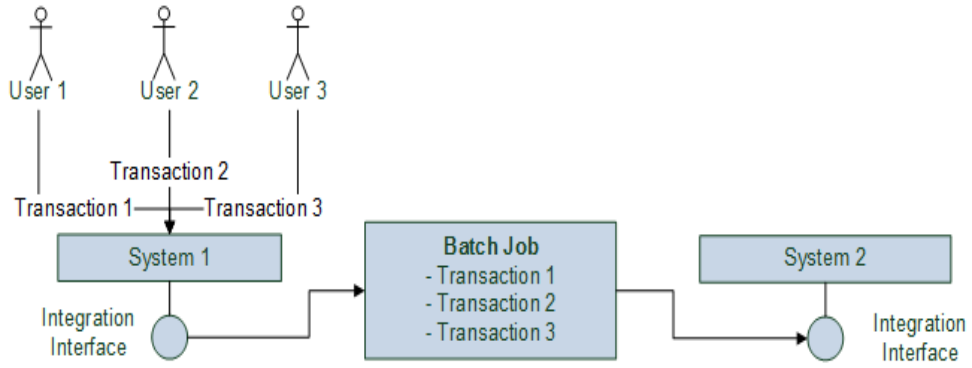


Figure 5. Batch Processing

In integration frameworks based on batch processing, a large group of transactions are collected and then the whole amount of data is processed by a software agent during a single execution. Since this usually involves a large volume of data processing tasks, it is better to

run such processes in time intervals of lower system resource load, for example during nights for systems which are loaded on the normal working hours.

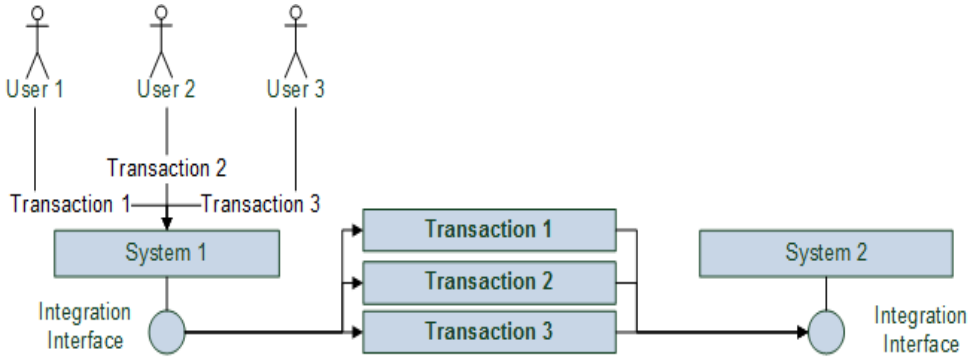


Figure 6. Real Time Processing

On the other hand, in a real-time processing based integration middleware, a single transaction or a small amount of data are processed on demand. Instead of a synchronizing the whole amount of data in a single run, the transactions are sent practically immediately to a remote software agent for processing. This guarantees that the two integrated systems are always synchronized and if a new event occurs against one of the systems it will be replicated against the other one in a real-time manner. Using this data processing approach secures that the data processing load will be spread across the whole time interval of system operation and not in a moment of single data synchronization execution.

4 Different Integration Approaches

Based on the above classification of business problems and definitions of different styles of data processing we can map them to different integration approaches [2]. Different integration styles are usually classified by the access of the integration interface to a layer of the application architecture [1]. Two of the most common approaches are the data integrationand the functional integration patterns.

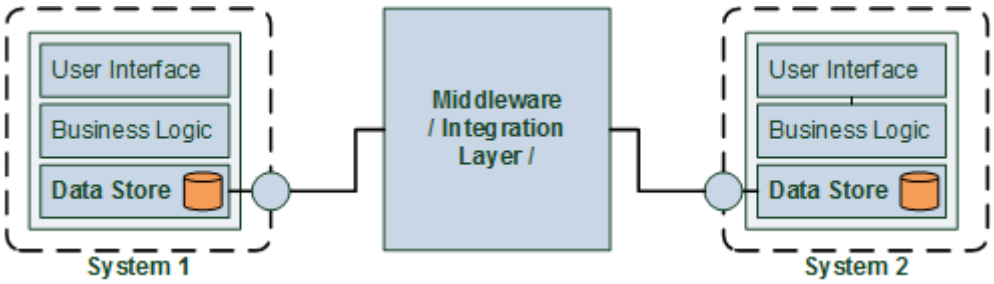


Figure 7. Data Layer Integration.

In data integration approaches, the integration adapter or interfaces are implemented to wrap directly the data repository or data-access layer of system that should be interconnected with another system or an integration environment [4]. Since this approach

by-passes the business logic layer and the already implemented business rules are not involved in the transactions processing, it is usually used when the main aim of the integration is synchronizing or replicating data between the repositories of the interconnected systems and direct access to data is sufficient.

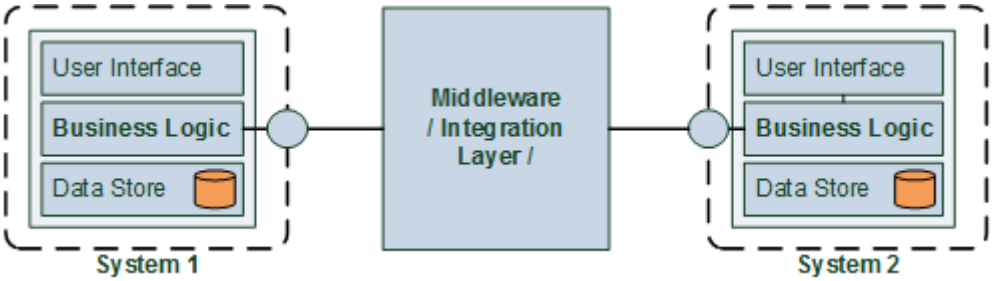


Figure 8. Functional Layer Integration.

In functional integration environments, the integration interfaces are implemented with wrapping the business logic layer of the stand-alone systems [3]. This can be done by invoking an existing API or if access to the system implementation is available. Such integration approaches are usually used, when direct access to data repositories is not appropriate and the existing business rules, like input validations, pre-save processing and so on must be explicitly used. The functional integration is often recognized as the most adopted and wide spread integration approach.

5 Mapping Between Different Problems and Integration Approaches

After knowing the basic classifications of business problems and integration approaches, we can map the both sets and research which mappings are more appropriate than others. For this aim we can use four models: real-time functional integration, batch processing functional integration, real-time data integration and batch processing data integration.

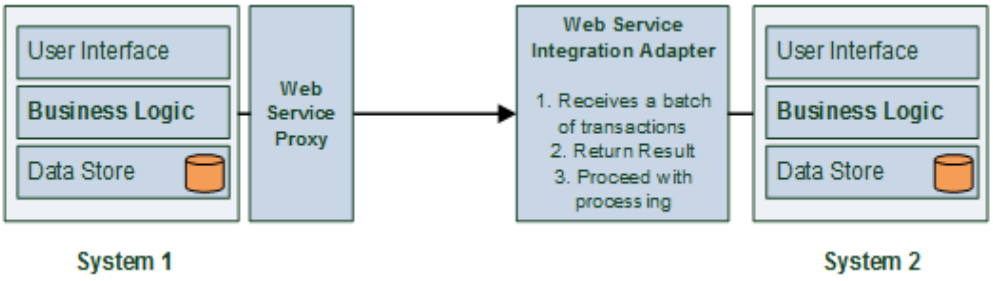


Figure 9. Functional integration model for batch processing.

In a sample functional integration model for batch processing, the target system has an integration interface implemented as a web service and the initiating system has a proxy which is based on the target system integration service. In this type of integration, the

initiating system calls the remote web service when the whole transaction batch should be processed, for example on scheduled intervals. The whole batch of transactions is sent in the service request, and then the web service iterates the transactions and calls the wrapped business logic in order to process them. After the processing is completed, the web service returns an asynchronous response to the calling system, which is not blocked during the batch processing.

Since the integration adapter is only responsible for iterating the batched transactions, its implementation is easier and this is the main advantage of this approach.

But while the wrapper implementation is easy, the method has significant disadvantages, for example that calibrating an integration middleware is not a simple task, which makes the integration between systems harder. The method works well only for small data volumes and can't be controlled synchronously, which is a serious drawback for functional integrations.

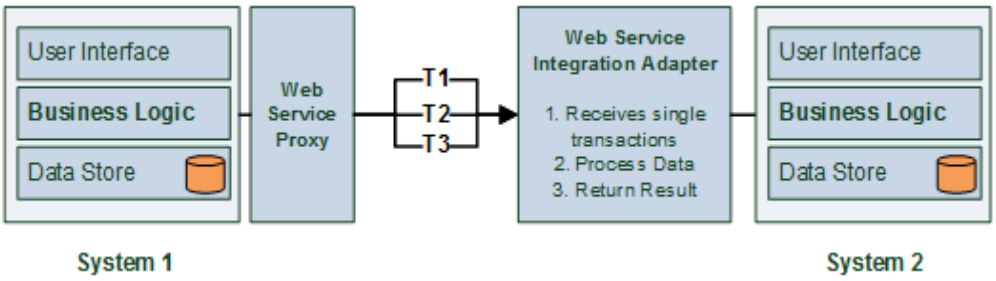


Figure 10. Functional integration model for real-time processing.

In an example of functional integration scenario with real-time processing, the same integration adapter setup can be used – a web service adapter wrapping the target system business logic layer and a web service proxy on the initiating side. The difference from the previous model is that the initiating system calls the target web service every time when a transaction event occurs. The web service request contains only single transaction data and the target system process the data in a real-time manner. After the processing is completed, the target system returns synchronous result to the initiating side, which is blocked during the processing period.

This method has some significant advantages. Since the integration adapter operates in a real time manner, the integration itself is easier, every transaction is processed on demand and the data volume is theoretically unlimited, because the processing load is distributed along time. The remote calls can be synchronously controlled also, which is an important advantage in functional integration designs.

The main disadvantage of this method is the harder integration adapter implementation, which is related to on-demand processing, synchronous execution and session and context management.

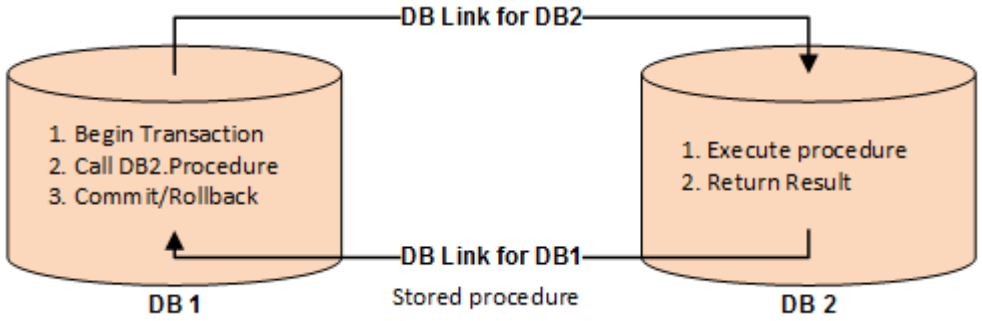


Figure 11. Data integration model for on-line processing.

Another sample integration framework is data integration for real-time processing. In this model the initiating system begins a transaction and delivers data to the connected system as parameters of a stored procedure, which belongs to the target system and is accessible via DB link. The target system completely performs the processing and then returns some result to the source system. If no exception had occurred, the system which initiated the transaction commits work, otherwise rolls it back.

The main advantages of this approach is the ease of the implementation of a single transaction processing stored procedure, used as an integration adapter, and the availability of synchronous control.

On the other hand, the method has significant drawbacks, like the large amount of stored procedure calls between databases, and the higher rate of network DBMS processing overload. These disadvantages make it inappropriate for large data volumes.

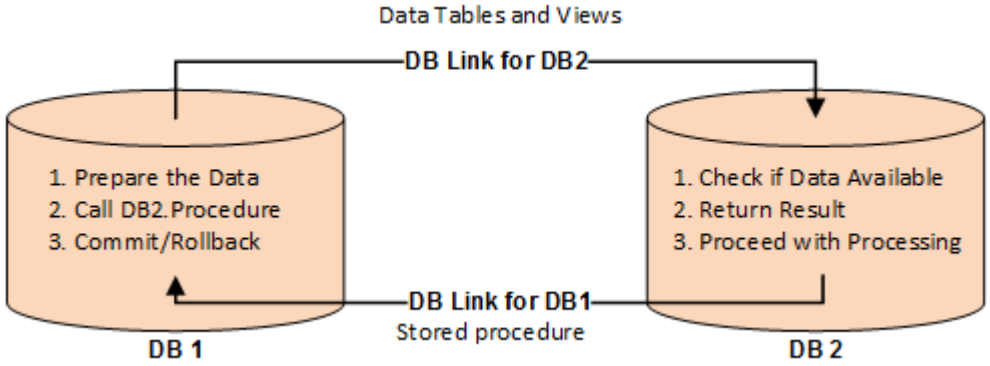


Figure 12. Data integration model for batch processing.

The last sample approach is a data integration framework for batch processing. In the proposed model the initiating system prepares all the data in suitable structures (tables and/or views) which are accessible to the target system via DB link. Then the source system begins a transaction and invokes a stored procedure, which belongs to the target system and is accessible via DB link. The target system performs some basic validations (for example checks if needed data are available) and then returns some result to the source system. If no exception had occurred, the source system commits work, otherwise rolls it back, and the target system proceeds with the data processing. That way two of the basic requirements are

satisfied – the transaction is as short as possible, and some quality of the service is achieved, because at least the successful startup and the most important prerequisites of the synchronization are guaranteed.

This approach involves a single stored procedure call and since this leads to low network and DBMS overload, it can be defined as a serious advantage. Based on these characteristics we can make the conclusion that the method is better for large data volumes.

On the other hand, implementation of transaction iterating and batch processing stored procedures is harder due to the query orientation of SQL. The stored procedure has limited synchronous control also and these limitations can be qualified as drawbacks of the approach.

6 Conclusion

After we have identified some different business problems and needs, we made a quick comparison of real time and batch processing and we looked through different integration approaches. Using the mapping between different business problems and integration approaches we proposed a couple of integration scenario samples and researched the advantages and disadvantages of various integration framework types.

Based on our comparison and analysis, we can conclude that data integration frameworks are better applied in business scopes with batch oriented problems, because of their lower network and processing overload and better scalability for large data volumes. On the other hand, functional integration frameworks are well adopted in real time processing problems, because of their better synchronous control, good data validation and convenient implementation of the integration middleware and respectively easier integration between systems.

References

1. Eric Long, *When Should You Use Real-Time Data Integration*, March 2013, jumpmind.com
2. John Grundy, Jun Bai, John Blackham, John Hosking and Robert Amor, *An architecture for efficient, flexible enterprise system integration*, Department of Computer Science, Department of Electrical and Electronic Engineering, University of Auckland, XSol Ltd
3. Krishnakumar Balasubramanian, Douglas C. Schmidt, Zoltán Molnár, Ákos Lédeczi, *System Integration using Model-Driven Engineering*, Institute for Software Integrated Systems, Vanderbilt University, Nashville
4. Tom Haughey, *Data Integration and Sharing - Part One*, October 1, 2003, TDAN.com
5. S. Kouzmanov, M. Tsaneva, *Structured Integration Approach Based On Levels Of Reusability*, International Conference On Application Of Information And Communication Technology And Statistics In Economy And Education (Icaictsee –2012), October 5-6th,2012, UNWE, Sofia, Bulgaria, pp.257-263.

Automated Verification of Machine Learning-Induced Knowledge Using Document Repositories

Milan Zorman[1,2], Sandi Pohorec[1], Bojan Butolen[1], Bojan Žlatanic[1], Dejan Dinevski [3], Peter Kokol[1]

[1]University of Maribor, Faculty of Electrical Engineering and Computer Science, Smetanova ulica 17, SI-2000 Maribor, Slovenia

[2]Centre for Interdisciplinary and Multidisciplinary Research and Studies at the University of Maribor, Krekova ulica 2, SI-2000 Maribor, Slovenia

[3]University of Maribor, Faculty of Medicine, Slomškov trg 15, SI-2000 Maribor, Slovenia

Abstract. Are times, when all the important knowledge was centralised only in paper form, becoming a matter of the past? An information era is to blame that more and more new and potentially new knowledge is primarily accessible in the electronic form. We cannot say that the written knowledge is not useful or obsolete, but not using or even ignoring the knowledge, accessible only in electronic form is by no means a smart decision. So, how to embrace the challenge and start using the available knowledge? In this paper we present an approach, which helps us to combine medical knowledge from two sources: symbolic machine learning approaches and classical data mining, and; natural language processing of electronically accessible texts. The knowledge from the first source is induced from a set of cases (tacit and implicit knowledge), provided by human experts and the second source are more conventional natural language texts, like papers, reports, e-books, and other explicit forms of knowledge from large repositories.

Keywords: machine learning; automated verification; rules; data mining; natural language processing; discourse analysis

Introduction

Modern medicine is an exact science based on evidence. Progress is ensured by on-going research throughout all areas. Over the past two centuries research and clinical trials have resulted in vast improvements in the treatment of diseases. In contrast to18th century medicine, modern medicine is evidence based [1]. Decisions about patients are made by usages of the best evidence available. Traditionally knowledge was accumulated from printed books. Over more recent years however the exponential growth of medical literature and research results has resulted in online publishing. Online papers can also be used for automated verification of results obtained from machine learning approaches. Similarly to the explosive growth of published papers with latest research results, machine learning also produces vast potential quantities for acquiring new knowledge.

Machine Learning-Induced Medical Knowledge

Machine learning (ML) knowledge induction processes provide us with two different forms of knowledge acquisition: non-symbolic that is rather difficult for humans to understand by humans and symbolic, which is closer to human reasoning and thus more appropriate for different kinds of usages such as classification, prediction, and knowledge mining.

Some domains, that also include medicine and biomedicine, often require detailed insights into the reasoning processes behind classifications, and the ability to understand relationships between attributes and final decisions, so symbolic knowledge presents a logical choice. The selection of appropriate methods for analysing data is therefore of crucial importance for success – for the acceptance and application of the results.

Symbolic ML induction approaches are seemingly a logical choice for data mining within this domain because they usually manifest their powers in combining computational power with reasoning procedures, which can then be easily followed by humans. Those approaches that provide us with full insight into the process of reasoning are also called 'white box' approaches. They collate details from the databases of already solved cases, so that any acquired knowledge is actually evidence based. As a by-product to classifications and predictions, we can acquire knowledge in the forms of rules, decision trees, lists of attributes that are necessary for decision making, etc.

Symbolic ML induction approaches are capable of dealing with three kinds of knowledge: tacit, implicit and explicit [2]. In order to convert tacit and implicit knowledge into a more desirable form, the produced explicit knowledge has to be evaluated by either human experts or automatic verification approaches, as presented in this paper.

We have to determine which knowledge presented in the forms of rules or sets of attributes, and obtained through the knowledge acquisition phase, is actually useful.

In general we differentiate between three types of extracted knowledge:

- Meaningless knowledge, usually produced during the later stages of knowledge action, and evaluation, where algorithms try to make trade-offs and compromises
- Existing knowledge, which usually refers to the level of knowledge already known by the expert responsible for the acquired knowledge's evaluation
- Potential newly acquired knowledge, either unknown to the expert, or the most valuable of all types of knowledge, newly acquired knowledge.

During the knowledge evaluation process the most difficult task is to find the difference between the meaningless and the potential newly acquired knowledge. In order to make this process easier for the domain expert, we have introduced the automatic verification of ML-Induced medical knowledge using document repositories.

In the following sections we briefly describe some of the more frequently used symbolic ML knowledge induction approaches:

- **Decision trees** (DT) are a typical representative of a symbolic machine learning approach used for the classification of objects [3]. The structure is similar to a flow-chart tree structure, each internal (not a leaf) node is a test on an attribute value (each object in the dataset is represented in the form of an attribute-value vector). Branches represent outcomes of node tests, each leaf node is a decision on the class to which an individual belongs to (each leaf is a class label) [4]. Decision trees are used for classification in such a way that, given an object U for which the classification class is unknown, the attribute values of U are tested against the nodes in the tree. The object U follows a path through the tree to a leaf node that assigns it with a class label. The decision tree is a popular classifier because it allows exploratory knowledge discovery as it does not require any domain knowledge and setting of parameters. The first algorithm for DT induction ID3 was introduced by Quinlan [5], [6] in 1986. Quinlan also presented C4.5 [7], a benchmark for newer supervised learning algorithms.
- **Association rules** approach searches for frequently occurring relationships and correlated relationships amongst attributes as a form of 'attribute-value' conditions, describing a set of data items [8]. The information or acquired knowledge is produced in the form of "if-then" statements or rules. The 'if'' part of the rule is called the antecedent and the 'then' part is called the consequent. The antecedent and consequent are sets of items that may not have any common items. The difference in comparison to previously presented decision trees is in the consequent, where decision trees use only one consequent attribute for all rules they produce on one data set.

- Ensemble methods in ML rest on the assumption that only the synergetic combination of different models can unleash their full potential [9], [10]. To overcome some of disadvantages and limitations of a single ML method, it is sometimes enough to use different models of the same machine learning approach, for example, using many different decision trees for the training subsets of the same data set. In other cases approach relies on combining different ML methods. Typically we can combine any machine learning methods in a form of an ensemble, but if we use symbolic machine learning methods, we satisfy the condition about possibilities of interpretation and explanation. The two more popular ensemble techniques are called bagging [11] and boosting [12]. If we simplify, we are basically speaking about combining multiple classifiers within a voting body, capable of making decisions with higher accuracy than any single classifier included within the voting body.

Information retrieval from document repositories

Information retrieval (IR) is a scientific field concerned with algorithms and methods for automated retrieval of information from document repositories. When dealing with large repositories, heterogeneity is a very important aspect to consider. Large repositories are inherently composed of documents with different structures (in terms of paragraphs and section annotations), length (sizes), distribution of topics changes (concentrated, uniformly scattered), speech manner, writing styles etc. In order to support verification of machine-induced knowledge acquisition, the document repositories have to be processed with respect to: syntactic parsing, part-of-speech analysis, discourse segmentation and formalisation of natural language.

Syntactic parsing

Document repositories can contain text in various formats: plaint text, structured by mark-up languages (HTML) or in proprietary formats (for instance PDF). As these documents have been prepared for human consumption they provide visual formatting without explicit annotations of words, sentences, and paragraphs. Therefore the division of texts into units called tokens is the first step in any IR application. Alphabetic languages, such as English, separate words using spaces. Usually IR applications require words, sentences, and paragraphs that are used as basic units of processing [13]. Since the rules of syntax in natural language are ambiguous [14], effective tokenisation has to incorporate a resolution of ambiguity at the semantic level. We have focused on the resolution of ambiguity regarding sentence-ending punctuations, because sentences are the smallest units of text that can offer enough information to verify a statement from machine-induced knowledge.

Part-of-speech analysis

Part-of-speech (POS) tagging is the automated process of assigning grammatical classes to words [15]. The process is quite complex since POS can be based on semantic, syntactic distributional ground or morphological grounds. Ideally the chosen tag should predict the behaviour of nearby words; therefore a semantic understanding of the text is required. POS tagging has been a very active research area for many years. The result of POS tagging is in the form of a parse tree (Fig. 1).

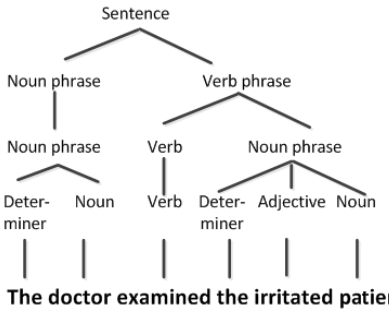


FIGURE 1: The parse tree of the parts-of-speech analysis for the sentence "The doctor examined the irritated patient".

A variety of implementations have been introduced over the years, the majority of them based on statistics. The TnT tagger is a good example of a relatively simple tagger that works fast and has an accuracy score of over 96% [16].

Discourse segmentation

If we consider visual formatting of documents we can conclude that they are segmented into logical areas. Scientific papers usually have the following structurally-defined units: title, abstract, introduction, multiple content chapters, conclusion, and references. These segments are usually further segmented into paragraphs that follow the discourse. A new paragraph implies a change in discourse or the focusing on a particular area. Automated discourse segmentation is required for a focused extraction of knowledge from documents. The process breaks documents into subject coherent multi-units (a unit being a sentence or paragraph). The basic idea is that the texts have to be searched for those parts where shifts of topics occur. These parts are then interpreted as the boundaries of the new topically-coherent segments. The task of segmentation is comprised of the following processes: division of the input text into smaller units, measurements of topic coherency between individual units and boundary selection.

The division of text can be implemented as a fixed-size window (x words, regardless of sentences) or various length (each unit is a sentence; the length is determined by the number of words in a sentence). Hearst suggested fixed lengths of 20 words as individual units [17]. This division alleviates the problems of punctuations creating multiple meanings, and the consequent splitting of text into sentences.

Measurements of topic coherency measure the amount of topic continuity. The measurement is done regarding the amount of evidence that the same topic prevails in both segments. The problems with coherency are that the coherency is relative to the position in the text (introductions are usually much more diverse in the topics they cover, main content chapters are longer and cover the same topic for multiple paragraphs/pages). Several methods of cohesion evaluation have been proposed [15]: *Vector Space Scoring* (basically evaluating the number of shared terms in the previous and next two units of text), *Block comparison* (only includes one text unit on either side and does not take inverse document frequency into account), and Vocabulary introduction (negative number of new terms in the left and right text units).

Formalisation of natural language

The formalisation of natural language is the transformation of the ambiguous notation of natural language into a formal notation with unambiguous interpretation of the meaning.

Traditional formal notations are semantic networks that represent relationships between individual words. A 'semantic network' is a term that encompasses a family of graph-based representations that share a common set of assumptions and concerns. A semantic network can be visualised as a graph in which node connections are the relationships between concepts. The concepts and types of association are labelled nodes. A well-used transformational practice is the mapping of nouns to concepts, and verbs to the connections between nouns. The problem of transformation can be easily viewed in the following interpretation: a sentence forms a vector of words (no punctuations) and the transformation process is tasked by assigning each word to its appropriate type (concept or link). As such, the transformation is a classification process: types of semantic network building blocks are the classes and the classification method assigns each word (or a multi-word sequence entity) to the appropriate type. The classification method uses POS tags of individual words along with semantic information for mapping to the appropriate type class. We present the process on a short sentence example:

Lithium doesn't slow Lou Gehrig's disease, first study on the drug stopped early.

Automated POS tagging returns the following vector of tags w_p = {[('Lithium', None), ('does', 'DOZ'), ('not', '*'), ('slow', 'JJ'), ('Lou', 'NP') , ('Gehrig', 'NP'), ("'s", None), ('disease', None), ('first', 'OD') , ('study', 'NN'), ('on', 'IN'), ('the', 'AT'), ('drug', None), ('stopped', 'VBD '), ('early', 'JJ'), ('.', '.')]}. The POS tags are those used in the Brown corpus [13].Semantic information helps resolve the words where POS couldn't be determined (Lithium is a silver-white metal, therefore it is a noun). The classification function Types (w_p) maps the sentence to the semantic net representation, as represented in Fig 2. The classification function is a supervised learning approach: we have manually annotated the more common sentence types found in scientific papers. Manual annotations were used as the training set for the naïve Bayes classifier, a simple probabilistic classifier based on the Bayes' theorem [18]. For each word, the conditional probability of it belonging to type T_i is calculated with the use of local word features f_i (within the current context). The chosen type is the highest ranked type according to formula 1, where *m* is the number of local word features (the features are assumed as being independent of the type). Probabilities P (T_i) and P ($f_j|T_i$) are occurrence frequencies in the training set for type T_i and feature f_j in the presence of type T_i.

$$T = \underset{T_i \in Types(w)}{\operatorname{argmax}} P(T_i|f_1, \dots f_m) = \underset{T_i \in Types(w)}{\operatorname{argmax}} \frac{P(f_1, \dots, f_m|T_i)\,P(T_i)}{P(f_1, \dots, f_m)} = \underset{T_i \in Types(w)}{\operatorname{argmax}} P(T_i) \prod_{j=1}^{m} P(f_j|T_i). \quad (1)$$

A semantic network is a graph-based representation with common assumptions and concerns. Visually it is represented as a graph where node connections are the relationships between concepts. Both nodes and connections are labelled in order to provide information about the represented concepts and association types. More recent models of representing natural language such as semantic networks are conceptual graphs [19](Fig 2).A conceptual graph is [**Error! Reference source not found.**] a finite, connected, bipartite graph. Conceptual graphs omit labelled arcs; instead relationship nodes represent the relationships between concepts. Conceptual graphs can be directly transformed into the triple form of ontology languages [**Error! Reference source not found.**]. Ontology form provides the capability of transforming the natural language text into unambiguous if-then type rules, useful for the final application (verification of machine learning results).

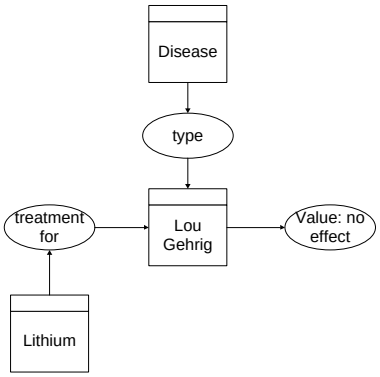


FIGURE 2: A semantic network representation of the sentence "Lithium doesn't slow Lou Gehrig's disease, first study on the drug stopped early."

Conclusion

The approach we have presented helps us to perform automated verification of machine learning-induced medical knowledge using document repositories by combining acquired medical knowledge from two sources: symbolic ML data mining knowledge, and natural language electronically accessible texts. In our tests we manually annotated the more common sentence types found in scientific papers in order to prepare the training set for the Naïve Bayes classifier, and build conceptual graphs. The conceptual graphs were then directly transformed into the triple form of ontology languages, which was used for reasoning and inference over the extracted formal data. This procedure provided the capability of transforming the natural language text into unambiguous if-then type rules, clearly comparable to induced ML knowledge.

Our approach presents a straightforward method of combining the results from two worlds and exploiting the vast resources of natural texts for automatic verification of ML-induced knowledge. Due to emerging ML approaches such as ensemble and other symbolic methods, the number of newly generated rules in every study grows and poses a problem for domain experts to thoroughly inspect and verify them.

Automated verification of acquired ML-induced medical knowledge using document repositories is therefore a useful tool, helping us to embrace the challenge and start using as much electronically- available sources as possible for acquiring medical knowledge.

References

1. Vrhovec L: Evidence based medicine. Acta Medico-Biotechnica. 2010; 3:11–8.
2. McGee J: Knowledge Management and Weblogs [Internet]. 2003 April [cited 09.05.2013], Available from: http://www.mcgeesmusings.net/2003/04/13.html#a3160
3. Podgorelec V., Zorman M. Decision Trees: Encyclopedia of complexity and systems science, New York: Springer. 2009; 2:1826-1845,
4. Han, J: Data Mining: Concepts and Techniques San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. ; 2005.
5. Quinlan JR. Induction of decision trees, Machine Learning. 1986; 1:81-106.
6. Quinlan JR. Simplifying decision trees, International Journal of Man-machine Studies.1987; 27:221-34.
7. Quinlan JR. C4.5: Programs for Machine Learning, Morgan Kaufmann, San Francisco, 1993.
8. Piatetsky-Shapiro G: Discovery, analysis, and presentation of strong rules. In: Knowledge Discovery in Databases (Piatetsky-Shapiro G, Frawley WJ, eds,). Cambridge, MA: AAAI/MIT Press. 1991;229 – 48.
9. Iglesias CJ: The Role of Hybrid Systems. Intelligent Data Management: The Case of Fuzzy/neural Hybrids, Control Engineering Practice. 1996; 4(6):839-45.
10. Wolpert D, Macready W: No Free Lunch Theorems for optimization. IEEE Transactions on Evolutionary Computation. 1997; 1(1): 7-82.
11. Breiman L.: Bagging predictors. Machine Learning. 1996; 24:123-40.
12. Freund Y, Schapire RE: Experiments with a new boosting algorithm. Machine Learning: Proc. Thirteenth International Conference. San Francisco: Morgan Kauffman, 1996; 148-56.
13. KuceraHM, Francis WN: Computational analysis of present-day American English, Brown University Press; 1967.
14. Grefenstette G, Tapanainen P: What is a word, what is a sentence? Problems of tokenization. In: Proceedings of the 3rd International Conference on Computer Lexicography. 1994; 79-87.
15. Manning CD, Schütze H: Foundations of Statistical Natural Language Processing, The MIT Press, Cambridge, Massachusetts; 1999.
16. Brants T: TnT - A Statistical Part-of-Speech Tagger, In: CoRR , Vol. cs.CL/0003055, 2000.
17. Hearst M: TextTiling: Segmenting Text into Multi-Paragraph Subtopic Passages, Computational Linguistics.1997; 23(1):33-64.
18. Navigli R: Word Sense Disambiguation: A Survey. ACM Comput. Surv. 2009; 41(2): 1-69.
19. Luger GF: Artificial intelligence, Structure and Strategies for Complex Problem Solving (Fifith Edition), Pearson Education Limited, USA; 2005.
20. Lacy LW: OWL: Representing Information Using the Web Ontology Language , Trafford Publishing , Victoria, CA; 2005.

Semantics Web and Ontology Learning

K. Musumbu

LaBRI (UMR 5800 du CNRS,
351, cours de la Lib´eration, F-33.405 TALENCE Cedex
musumbu@labri.fr

Abstract. The Semantics Web is a vision for the future of the Web in which information is given ex-plicit meaning, making it easier for machines to automatically process and integrate information available on the Web. An ontology defines the terms used to describe and represent an area of knowledge. Ontologies are used by people, databases, and applications that need to share domain information (a domain is just a specific subject area or area of knowledge, like medicine, tool manufacturing, real estate, automobile repair, financial management, etc). The amount of data caused by the success of Internet is demanding methodologies and tools to automatically extract unknown and potentially useful knowledge out of it, generating structured representations with that knowledge. The main goal of this paper is to present comprehensive introduction into MDA- based ontology development. It will provide an introduction to the field of the Semantics Web and ontology engineering, a description of several UML- and metamodeling- based solutions and tools for ontology development.

Keywords. Business rules, knowledge based systems, model driven architecture, reasoning, semantics web, ontology.

1. Introduction

Certain commonly used terms had multiple meanings all equally valid but which we had not differentiated adequately, leading to much confusion. To avoid this confusion one must define an ontology, which included semiformal definitions of all key terms and their relationships. The term ontology means a specification of a conceptualization. It is used to describe the high level structure of the software in the most direct and most expressive manner possible. The Semantic Web endows the network of disparate information that is currently accessible on the Internet with meaning. Because this meaning can be gleaned and processed automatically by software, the Semantic Web opens up the exciting and awe-inducing possibility of a unified global intelligence accessible to all. The e-learning is the use of the new multimedia technologies of the Internet to improve the quality of the learning by facilitating on one hand the access to resources and to services, on the other hand the exchanges and the collaboration in distance. The main goal of this paper is to present comprehensive introduction into MDA-based ontology development. It will provide an introduction to the field of the Semantics Web and ontology engineering, a description of several UML- and metamodeling based solutions and tools for ontology development. The object if is to show how the ontologies can be useful in the domain of e-learning. This paper will be divided in three parts. The first covers the basis of both main topics ontology, semantic web and standards. The second part, explain the Model Driven Architecture. The last is the central one, it starts with a review of several approaches and aim to bridge the gap between ontology development and software engineering methodologies.

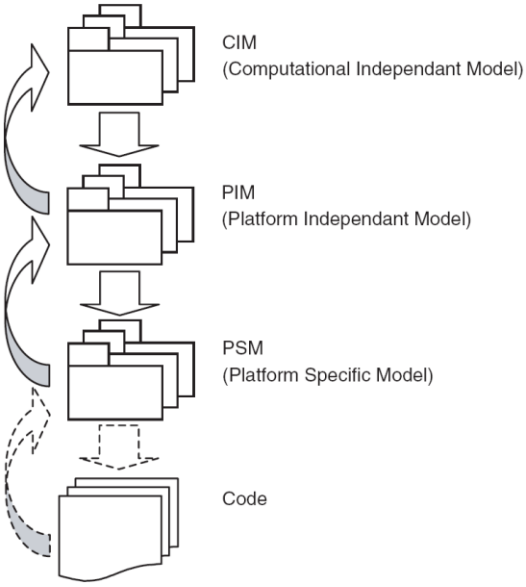


Figure 1: Global view of the Model Driven Architecture approach

2. Semantic Web

The Semantic Web has the potential for semantically richer representations of things (e.g., Web pages, applications, and persons) and their relations on the Web, and thus should provide us with more intelligent services. In MDA, an instance of MOF [10] is use for representing models but our works are only concerned by UML models. For adding semantics in UML models we can use:

- *UML profile*: UML can be used for modeling many domains. The problem with this is that UML models are so generic that it is impossible to know either it is object application, a metamodel, a model, a database structures or anything else only by looking at it [9]. For adding precision, the OMG has standardized the concept of UML profile [11]. A UML profile is a set of techniques and mechanisms allowing to adapt UML in a particular and specific domain. UML profile can be used in any UML model and do not modify the structure of the meta-model. UML profiles are stereotypes or labels which can be injected into models. After having stuck labels on models, we can make inference using it. As we can see, doing this can solve our problem of semantics lack on model in a low level, but this is not exploitable by machines because there is no notion of logic and taxonomy and semantic is not formally defined.
- *Object Constraint Language*: In UML it was not possible to define the body of an operation (or a method) so the OCL [12] was standardized by OMG for doing it. OCL allows expressing any kind of constraints on UML models. For example, we can express constraints like: “before renting a car you must be sure that it is OK”. Well, OCL seems to be a good solution for our problem but it is not the case. The first

problem with OCL is that he does not support side effect operations and the second is that he does not offer automatic inference for machines.

- *Action Semantic*: remember that the main constraint with OCL was that he only supports no side effects operations. To solve this constraint, the OMG standardize Action Semantic [13]. Well, now we have a formalism being able to express any king of operations and constraints but it is not enough. This formalism is complicated to use [9], was not created while thinking to machine comprehension and self-use, and do not have a textual formalism.

As we can see, none of the UML "techniques" is suitable for adding semantics in models. In another side a new domain of computer is growing more and more: semantic web. The aim of the semantic web is to make the web both comprehensible by humans and machines [14]. A part of semantic web is about ontology and reasoning. Modeling concept defined by ontologies can be used to model the concepts in a domain, the relationships between them, and the properties that can be used to describe instances of those concepts [15]. In addition, the Web Ontology Language (OWL)[24] supports the inclusion of certain types of constraint in ontology, allowing new information to be deduced when combining instance data with these logic's description [15]. At this point our dilemma was how can we use MDA models and Semantic Web? Ontology Definition Meta Model was the response to our need.

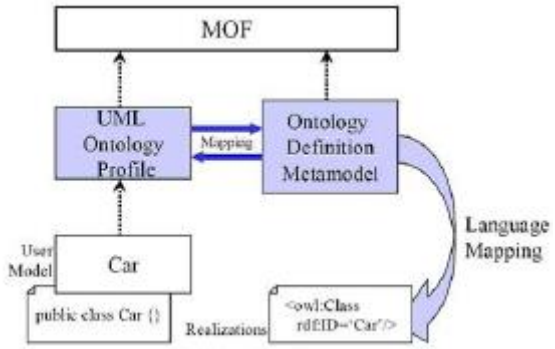


Figure 2: ODM principe

Moreover, careful selection of a knowledge representation may simplify problem solving, whereas an unfortunate selection may lead to difficulties or even failure to find a solution. Complex problems require a combination of several different representations. In order to be practical, every knowledge representation technique needs a notation. A formal notation often used for representing knowledge is first-order predicate calculus. Of course, it is not suitable for all types of knowledge. The notation of description logics has also proven expressive enough to represent knowledge formally. More recently, the trend has been to use XML syntax for representing knowledge structures.

3. Ontologies and their application

There is an increasing awareness among researchers and developers that ontologies are not just for knowledge-based systems, but for all software systems. The major application

fields for ontologies nowadays include knowledge management, e-learning, ecommerce, and integration of Web resources, intranet documents, and databases. Knowledge sharing and reuse through automatic exchange of Web documents among applications and agents is possible only if the documents contain ontologically encoded information. One view of ontology representation languages is that they can be used as metadata languages ontologies engineered with such languages can provide vocabularies for metadata to facilitate the management, discovery, and retrieval of resources on the Web. A metamodel is an explicit model of the constructs and rules needed to build specific models within a domain of interest. This characterizes a valid metamodel as an ontology, since such constructs and rules represent entities in a domain and their relationships i.e., a metamodel is an ontology used by modelers.

The MDA and its four-layer architecture provide a solid basis for defining the metamodels of any modeling language, and thus a language for modeling ontologies based on the MOF. Such a language can employ MDA tool support for modeling, model management, and interoperability with other MOF-based metamodels. The current software tools do not implement many of the fundamental MDA concepts. However, we can expect that most of these tools, which are presently oriented towards UML and the modeling layer (M1), will be improved and equipped with MDA support in the coming years.

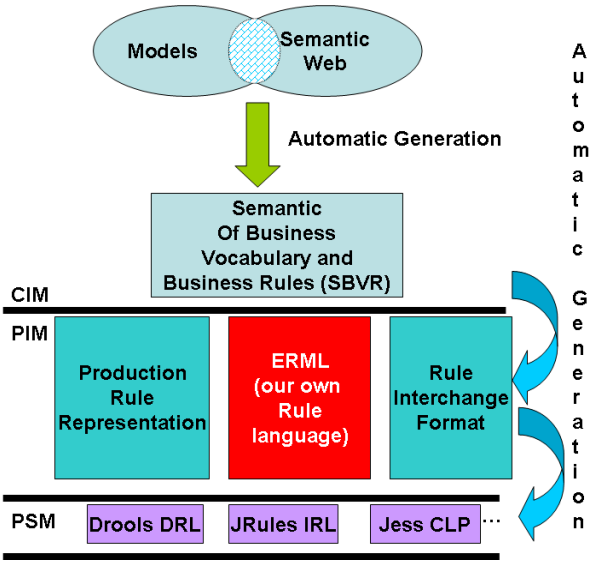


Figure 3: Our approach throughout the MDA layers

4. Ontologies Learning

Ontology learning (ontology extraction, ontology generation, or ontology acquisition) is the automatic or semi-automatic creation of ontologies, including extracting the corresponding domain's terms and the relationships between those concepts from a corpus of natural language text, and encoding them with an ontology language for easy retrieval. As

building ontologies manually is extremely labor-intensive and time consuming, there is great motivation to automate the process. Typically, the process starts by extracting terms and concepts or noun phrases from plain text using linguistic processors such as part-of-speech tagging and phrase chunking. Then statistical or symbolic techniques are used to extract relation signatures. ontologies have to be structured optimally against the knowledge base they represent. The structure of an ontology needs to be changed continuously so that it is an accurate representation of the underlying domain. Recently, an automated method was introduced for engineering ontologies in life sciences such as Gene Ontology (GO),[7] one of the most successful and widely used biomedical ontology.[8] Based on information theory, it restructures ontologies so that the levels represent the desired specificity of the concepts. Similar information theoretic approaches have also been used for optimal partition of Gene Ontology.[9] Given the mathematical nature of such engineering algorithms, these optimizations can be automated to produce a principled and scalable architecture to restructure ontologies such as GO.

5. Conclusion

Ontology engineering aims to make explicit the knowledge contained within software applications, and within enterprises and business procedures for a particular domain. Ontology engineering offers a direction towards solving the interoperability problems brought about by semantic obstacles, such as the obstacles related to the definitions of business terms and software classes. However, real world applications and their development are rarely clearly bordered in their scope; that's why these environments are not enough. It is, therefore, necessary to integrate applications that are used for intelligent systems development into mainstream software platforms. This topic is going to gain more and more attention with the development of the Semantic Web [Berners-Lee et al., 2001] and with increased integration of intelligent techniques in common information systems.

References

1. Barbara von Halle. *Business Rules Applied*. John Wiley & Sons, New York, USA, 2002.
2. Ronald G. Ross. *Principles of the Business Rule Approach*. Addison-Wesley, Boston, USA, 2003.
3. The Object Management Group OMG. *Semantics of Business Vocabulary and Business Rules (SBVR)*. OMG Specification, March 2006.
4. W3C. *Rule Interchange Format (RIF)*. W3C Workgroup, 2005.
5. The Object Management Group OMG. *Production Rule Representation (PRR) RFP*. OMG Request For Proposal (br/2003-09-03), 2003.
6. IBM T.J. Watson Research Center. *Common Rules project*. Intelligent Agents project (1994-97), 1997.
7. RuleML. *The Rule ML initiative.*
8. The Object Management Group OMG. *Model Driven Architecture Guide Version 1.0.1*. OMG Specification, June 2003.
9. Xavier Blanc. *MDA en action*. Eyrolles, France, 2005.
10. The Object Management Group OMG. *Meta Objet Facility (MOF) Specification Version 1.4*. OMG Specification (formal/02-04-03), April 2002.
11. The Object Management Group. *Unified Modeling Language: Superstructure*. OMG Specification, February 2004.
12. The Object Management Group OMG. *UML 2.0 OCL Specification*. OMG Specification, October 2003.
13. The Action Semantics Consortium. *Action semantics for the uml*. OMG Specification (ad/2001-03-01), March 2001.

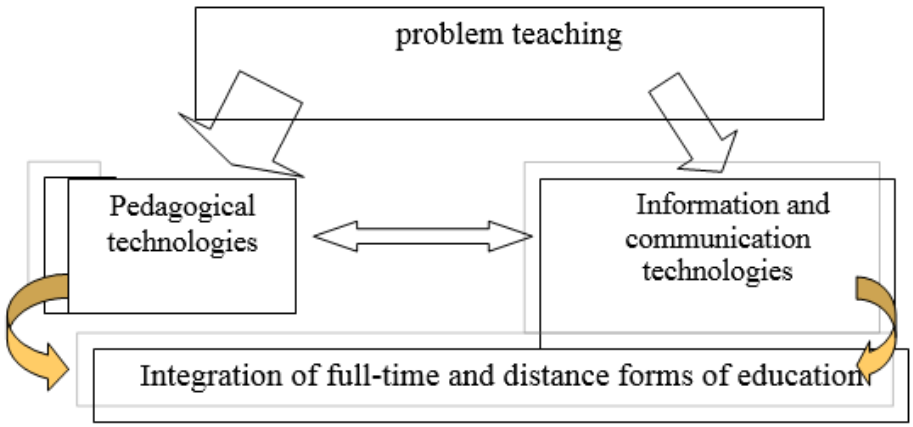
14. Thomas B. Passin. *Explorer's guide to the Semantic Web*. Manning Publications Co, Greenwich, UK, 2004.
15. Stephen Cranefield and Jin Pan. *Bridging the Gap Between the Model-Driven Architecture and Ontology Engineering*. *Proc. of AOSE 2004 Workshop*, 2004.
16. Dragan Ga˘evi`c, Dragan Djuri´e, and Vladan Deved˘zi´c. *Model Driven Architecture and Ontology Development*. Springer-Verlag, Berlin, DE, 2006.
17. Gruber, Thomas. Toward Principles for the Design of Ontologies Used for Knowledge Shar ing. International Journal Human-Computer Studies Vol. 43, Issues 5-6, Nove mer 1995.
18. The Object Management Group OMG. *Request For Proposal for Ontology Definition Metamodel*. OMG Request For Proposal, March 2003.
19. Kenneth Baclawski, Mieczyslaw K. Kokar, Paul A. Kogut, Lewis Hart, Jeffrey Smith, William S. Holmes III, Jerzy Letkowski, and Michael L. Aronson. *Extending UML to Support Ontology Engineering for the Semantic Web*. *Lecture Notes in Computer Science*, 2185:342+, 2001.
20. Dragan Djuric, Dragan Gasevic, and Vladan Devedzic. *Ontology Modeling and MDA. Journal of Object Technology*, 4(1):109–128, 2005.
21. The Object Management Group OMG, IBM, and Sandpiper Software. *Ontology Definition Metamodel*. OMG Specification, June 2006.
22. H. Knublauch. *Ontology-Driven Software Development in the Context of the Semantic Web: An Example Scenario with Protege/OWL*. *1st International Workshop on the Model-Driven Semantic Web (MDSW2004)*, 2004.
23. Harold Boley, Said Tabet, and GerdWagner. *Design Rationale for RuleML: A Markup Language for Semantic Web Rules.* In *SWWS*, pages 381–401, 2001.
24. Stephen J. Mellor, Anthony N. Clark, and Takao Futagami. *Guest Editors' Introduction: Model-Driven Development. IEEE Software*, 20(5):14–18, 2003.
25. Raphael Volz. *Web Ontology Reasoning with Logic Databases*. PhD thesis, Universit¨at Karlsruhe (TH), Universit¨at Karlsruhe (TH), Institut AIFB, D-76128 Karlsruhe, 2004.
26. The Object Management Group OMG. *MOF 2.0/XMI Mapping Specification, v2.1*. OMG Specification (formal/05-09-01), 2005.
27. Java Community Process(JCP). *Java Metadata Interface (JMI)*. Sun Java Specification Request (JSR 40), 2002.
28. Mouhamed Diouf, Kaninda Musumbu, and Sofian Maabout. *Standard Business Rules Language: why and how? The 2006 International Conference on Artificial Intelligence*, June 2006.

Applying of Modern Information and Communication Technologies in Education

Mamedov H., Babanli M., Jafarov Z.
Azerbaijan Technical university, Baku, Azerbaijan
babanli@aztu.edu.az

Abstract. The information and communication technology, used in educational process is analyzed from several aspects. It is proved, that with application of information and communication technology it is possible to create effective system for remote and continuous education.

Modern progress level of development of information and telecommunication technologies focuses our attention on development of concept and structural and organizational models, their applying in informatization of education (picture 1).



Picture1. Modern paradigm of education

Pedagogical technologies:

- ✓ education in collaboration
- ✓ «brainstorming»
- ✓ situational analysis
- ✓ method of projects
- ✓ « student`s briefcase »
- ✓ network lecture
- ✓ network seminar
- ✓ network debates (discussions)
- ✓ tutorials
- ✓ network practical lessons
- ✓ network practical works
- ✓ network debates

- ✓ network consultations
- ✓ TV testing
- ✓ online games with education directionality
- ✓ virtual excursions

Information and telecommunication technologies:

1. Computer technologies based on CD-ROM:
   - text editors,
   - hypertext editors,
   - editors of multimedia presentations,
   - technology of virtual worlds (3-D).

2. . Web 1.0 technologies (internet technologies of first generation):
   - e-mail,
   - forum,
   - audio and video chat,
   - facilities of IP-telephony,
   - programs-shells (platforms) for creation of network courses (CMS) and organization of distance education (LMS).

3. Web 2.0 technologies (internet technology of the second generation):

- blogs,
  - Wiki (technology of collective hypertext),
  - services of social bookmarks,
  - services for storing and creating of Power Point presentations,
  - microblogs (services for professional communication),
  - social media storage (photo services, video services, podcasts),
  - contact services,
  - social and search systems.

4. Web 3.0 technologies (mobile technologies):
   - Smart-technologies
   - distributed computing,
   - mashup (allows you to create content)
   - Web-Syndication (RSS-Really Simple Syndication),
   - services for the creation of virtual worlds (Second Life - service for creation of three-dimensional virtual world)
   - «Cloud" technologies (Cloud computing),
   - Telepresence technologies, and etc.

Specifics of new pedagogical reality:

A new concept of education - continuing education throughout the life as a condition for a successful life;

- changing the educational paradigm;
- changes of students' attitudes toward the working with information;
- Formation of new direction of philosophy of education - connectionism (Siemens);

• Information and didactic space of educational process,

• A new type of learner – ICT literate user of educational services («digital natives») (Mark Prensky),

• changing of requirements to "professiograme" of teacher («digital immigrants»)

The main tasks for integration of continuous , full-time and distance education are:

- development of information infrastructure ;
- organizing of network interaction of educational institutions ;
- creation and development of digital educational resources, including taking into account national and regional component;
- improving of training system of pedagogical , managerial, engineering and technical personnel in the field of information technologies;
- ensuring of equal opportunities for all citizens (including people with limited health abilities) to receive education of all levels.

Justification of urgent need for introduction of computerized and telecommunication technology in the educational process of higher education in practice has two major, closely related terms. First, incomparable technical and operational capabilities of computers carry didactic material, which incomparable with earlier applied technical learning tools and which can and should be implemented in the training and educational process. Secondly, the true efficiency of scientific and technological progress depends on the training of personnel at the level of current requirements. Therefore, the use of computerized and telecommunication technology in the educational process - a key component of preparing of students for future work life. It is not true to ignore the fact that for the majority of graduates of higher education institutions future profession will be mostly related with information and communication technology (ICT). Education becomes one of the sources of the most valuable strategic resources - human capital and knowledge that ultimately determines the overall level of society development. And informatization becomes the main accelerator of its development. Informatization of society is almost impossible without the computerization of the education system. The priority of this problem is increased by the fact that it is a fundamentally new.

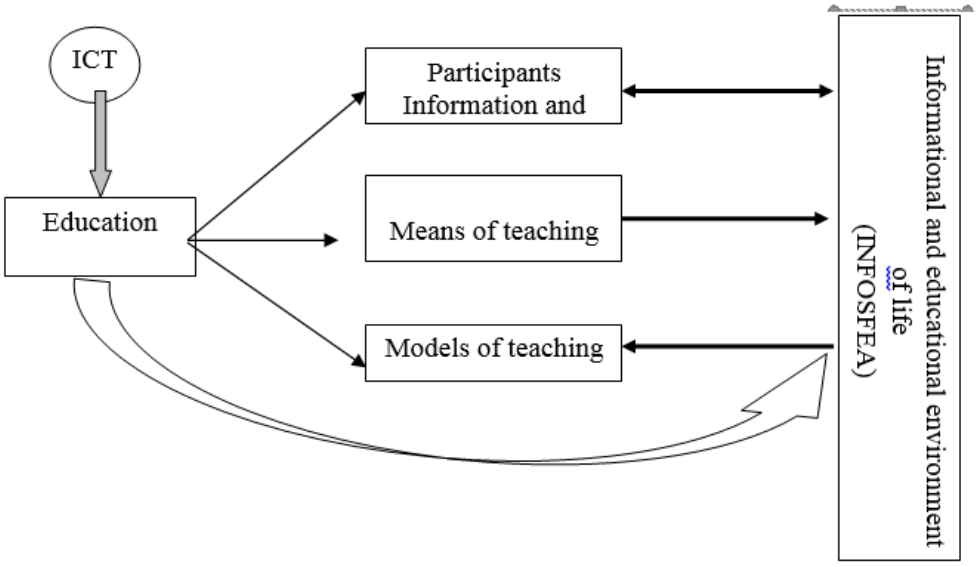
Applying of information and telecommunication technologies is the most important role in the new system of education. The main purpose of the use of ICT in education:

- improving of administrative performance ;
- support of managerial and scientific researches;
- expand of scopes of the learning process (the creation of open educational space , virtual education institutions) ;
- increase of efficiency of students independent activity, the use of specific properties of PC as a means of learning.

Comprehensive informatization of higher education institutions should focus on the formation and development of intellectual potential of science, improving of forms and content of educational process, the introduction of computer training methods , the use of modern information and communication technologies in the pedagogical work (Picture 2).

Development of education technology with using of technical means became sufficient with the presence of computers only. Previously performed functions: TV, VCR, tape recorder, movie projector, slide projector, etc., successfully took the information and communication technologies. And the quality of the transmission, storage, displaying of

information significantly improved. In this way, it is necessary to consider the task of translating of all information in the digital standards as a priority . ICT is a powerful tool for processing the information provided in the form of words, numbers , images, sounds , etc. The main feature of ICT as a tool is the ability of it's configuring (programming) to perform various activities related with receiving and recycling large amount of information.



Picture 2. Comprehensive informatisation of higher education institutions

The use of ICT in educational process opens new ways in the development of thinking skills and ability to solve difficult problems, provides a fundamentally new opportunities for accelerating of learning. ICT helps to do classroom and independent lessons more interesting, dynamic and compelling, and a huge flow of information easily accessible for study.

The use of ICT makes possible to make learning process more active, give it character of search and research. Unlike the books , television and movies ICT provides opportunity of immediate response to the action of student, repetition , explanation of material for the weaker students, and transition to a more complicated and supercomplex material for the most prepared students . At the same time teaching in an individual order easily and naturally realized. There is no doubt, that benefits of ICT in many cases are not disputed. It not only saves students from routine work, but also allow them to do labor-intensive practical tasks using a variety of methods and complex analytical researches.

ICT opens new perspectives in the field of education. In the process of increasing of volume of knowledge and complication of methods of analysis, it becomes increasingly difficult to build training, adhering to the principle of passive listening to lectures and reading of texts. In the present social and economic conditions the possibility of obtaining of continuous education is very important. The task of training and retraining of qualification is equally important, and distance education is an irreplaceable mechanism for receiving of quality education. Distance learning refers to the method of delivery of educational material

(interaction) within the framework of distance education and teaching in the distance - independent work in any form of education.

We single out the main components of the new education environment :

physical (actually audience , its size , lighting, microclimate, capacity , location of workplaces , etc.) ;

informational (virtual) which includes :

- hardware (component) ;
- software (including network) ;
- meaningful (including network) ;

methodical, which helps to structure and actualize the course material (here structured blocks: precis and organizer, the system of texts for the current and total control and self-control, take a special place) ;

psychological, taking into account the type, specificity and intensity of relationship between teacher and students, and students with each other in the real the real conditions of computer class (of course, this component is present in a traditionally organized education process, but the use of ICT radically increases the load of all its members);

intellectual, which are the least formalizable.

The first and main task of any IT – system, which is in the field of education, is a decrease of paper documents. The second task is often taken into account is the optimization of access to information. Examples of foreign universities give us the third task- facilitation of some university services. The fourth task it is a possibility of methodological support of network learning process.

There are a lot of different approaches of classification models of learning using the Internet, but in all cases, didactic and organizational basis is network educational - methodical and information complex (NEMIC).

The research of experience in composition teaching methods shows that the rational structure of the complex network of training includes the following basic elements:

- educational programs;
- teaching materials (teaching aid, lectures, presentations and so on);
- guide for self-study;
- reading book;
- control unit (tests, seminars, practical tasks, projects ...);
- glossary;
- a list of abbreviations and acronyms

Designing of NEMIC can be performed in the following sequence. First of all, we formulate goals and then we held the selection of content. Teacher`s work results is the training materials, issued in the form of a teaching aid that can be physically placed on CD or online.

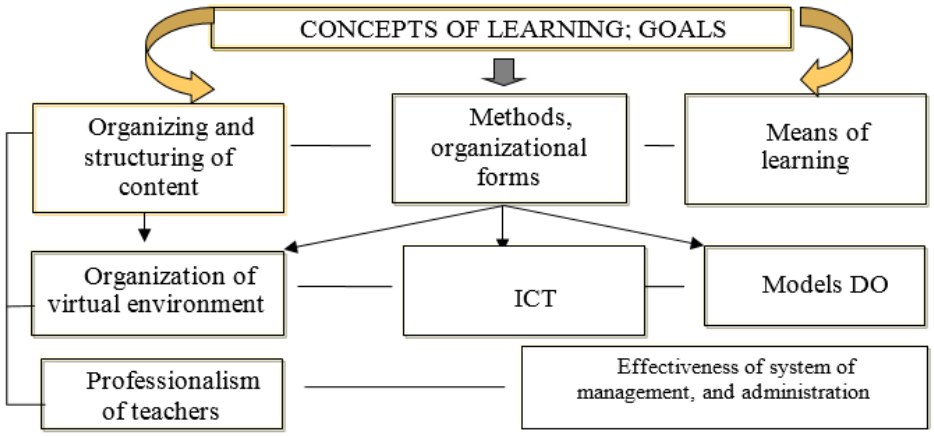
The process of formation of a unified information space of an educational institution became complete without providing organizational and technical conditions for the development of network interactions of educational institution: in addition to the creation of the corporate network space is becoming increasingly important task of information integration of the educational institution with external environment. We need to make our educational institution as an active participant in the Internet space. Besides passive usage of

the Internet as an additional universal source of information, we should develop such directions of the Internet as:

- activation of contacts with the environment;
- network support of internal training;
- development of representation of educational institutions in the external environment.
- distance education

Distance learning / e-Learning – it is a system of learning based on the interaction of subjects of the educational process between themselves and with the distance learning facilities using special information and communications educational environment.

The continuous form of learning is being made cheaper by ICT and encourages students to increase their knowledge on various subjects. Distance education (DE) is a real opportunity to get a quality education without actually living in the city where the person is going to learn (picture 3).The access to the informational resources of the libraries – it is also a problem which can be solved with the help of ICT. The usage of these technologies helps the student to form the new organization of work.



Picture 3. Factors affecting to the quality of distance learning

If at the traditional approach students listen lectures, conduct notes, visit libraries and seminars, it is actually built into an organized learning process. In the situation of distance education student must organize myself and get the necessary level of knowledge that can be tested using the test system.

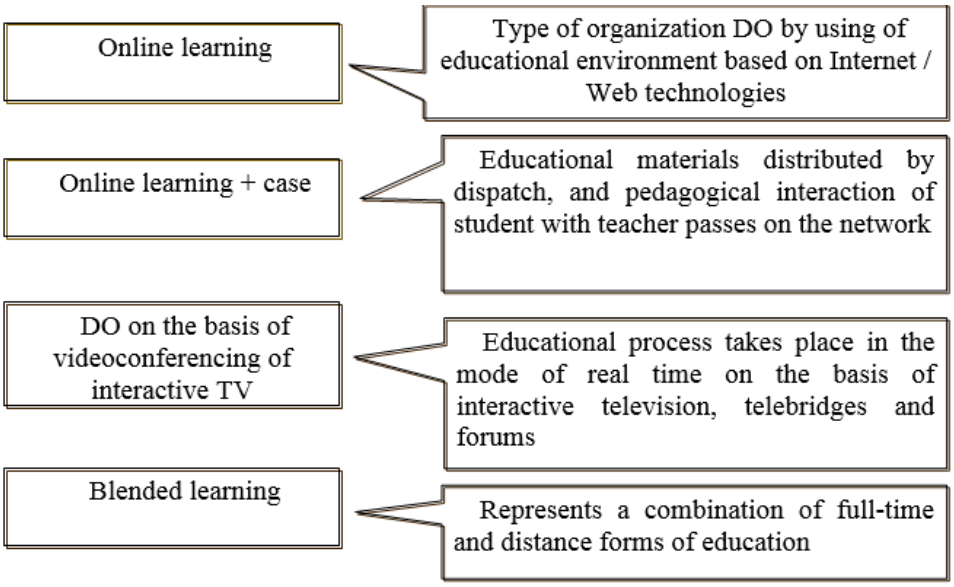
Current economic conditions of life practically in every country of world characterized by the need for a significant portion of students of all forms of training (day, evening and part-time) to combine study with work. In addition, some of the people who want to pursue higher education, deprived a varying extent the opportunities to study in high school of their choice they need to get a specialty. Visiting by student at the educational institution, as a rule, is episodic and is defined by the need to perform those functions which,

for whatever reasons, otherwise cannot be made by them. Necessary educational materials in printed, video, audio, and / or machine-readable form, at the request of the student shall be issued to him by hand or sent, for example, via e-mail. It is obvious that the use of this form of training became possible, first of all, thanks to the modern achievements of science and technology (in our case - the new computer and communication technologies in education). In addition, well-organized distance learning can be not only more effective but also more comfortable for the student, and the funds received for students learning can be effectively used for the organization of modern teaching technologies. Undoubtedly, this form of education cause demand of use completely different modern teaching technologies. In our opinion, at the time of creation of such technologies the main attention should be paid to the development in students' logical thinking skills, ability to work with different types of information, to prepare and to make appropriate decisions. This becomes even more actual that the student learns individually and he has often no one to consult. In this connection, technology of training should take into account the ability to work of students in the information and telecommunications networks, both independently and as part of the so - called "virtual study groups." In this case, students will not only acquire the necessary specialty, but also get a wide range of practical skills to work with modern computer systems, office equipment means and means of communication, different sources and forms of information.

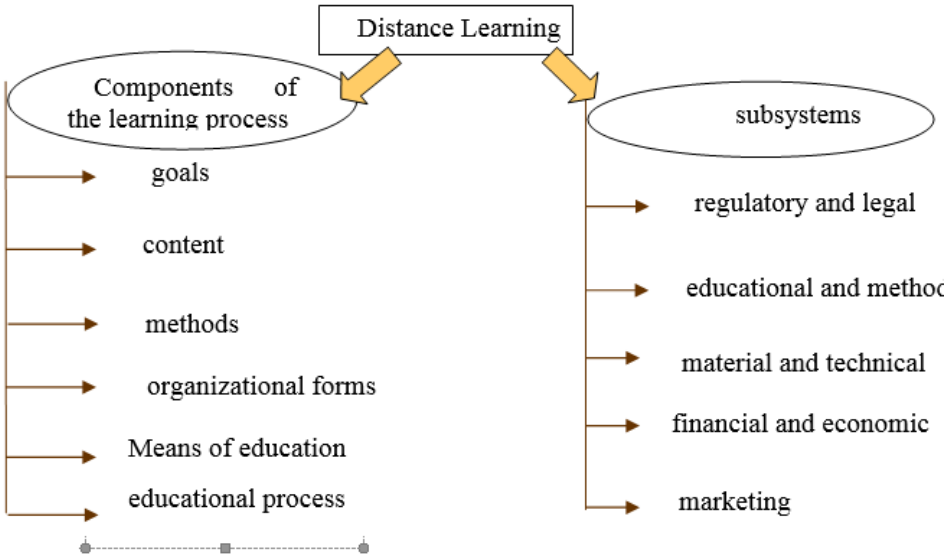
During the analysis, revealed that the most important organizational conditions of the distance learning system in a real educational practice include:

- Scientific development of theoretical frameworks and technologies DO;
- Design of model of organization of distance learning (Picture 4);
- Organization of approbation of this model and verification of its in mass teaching practice;
- Purposeful use of information and communication technologies in the process of distance learning;
- Creation and use of telecommunication complex in the system of distance learning system;
- Organization of design and development of didactic provision of distance learning process.

Means of distance learning include: multimedia technologies; electronic educational materials: textbooks , tests, trainers, problem books , manuals , etc.; virtual laboratories : network versions, CD-ROM, files; videoconferences, hypertext technology, teleconferences, distribution laboratories with remote access, electronic libraries (books , magazines , reference books , articles , etc.), virtual reality , Web- servers of the Internet , e-mail. Various forms of network communications (personal communication by e-mail, thematic lists, forums and chat rooms, methodical ftp- archives, etc.) can be considered as the basis for the formation of a integrated training and methodological space (Picture 5).



Picture 4. Models of distance learning



Picture 5. Formation of distance learning

One more possibility of using Internet technology is creating their own online courses, which should not cause any complications if the teacher is familiar with Internet technology and know how to use a simple editor for creating web pages.

Finally we can to make the following conclusions. Participation in the network of teaching unions helps teachers to exchange experiences, to discuss current problems with colleagues, get a "first hand" normative documents. Effective inclusion of educational

institutions in the external information space provides its information integration with other educational institutions, management bodies, contributes to the improvement of scientific and methodical work of teachers and, eventually, lead to qualitative changes in the educational and scientific process.

Thus, the inclusion of these means into practice will form the students' literacy skills to work with different kinds of information using information and communication technologies; to involve every student in an active learning process; realize free access to the necessary information not only in the national information centers of educational institutions, but also in the centers of other countries; to conduct communion with their peers from other educational institutions of the region, the regions of the country and even other world countries; to work in cooperation in solving various problems, showing in this time certain communication skills; to form ethic of work in telecommunication networks, etc. Advantages from the use of methods of elements of distance learning technologies: During teaching the student becomes a full member of the information society, because except of training of chosen specialty, gets skills of using modern technologies, which gives an advantage in subsequent employment and career growth.

References

1. ICT in Education of People with Special Needs**:** Specialized training course / Authorized Lane. from English. / N. Tokarev, S. Besio - M. Ed. House "Education Service", 2008. 320 p., 32 ill., 30 tab. ISBN 5290211622021

2. Information and Communication Technologies in Special Education**.** Final Report of the international expert meeting held in Moscow by IITE on12-13 April 2002.

Using Importance-Performance Analysis and Quality Function Deployment to Improve Medical Service Quality – A Case of Psychiatric Patients in Taiwan

Chiung-yu Yeh[1], Hsin-Hung Wu[2]

[1] Changhua Christian Medical Foundation Lutune Christian Hospital, Changhua, Taiwan
25219@cch.org.tw

[2] Department of Business Administration, National Changhua University of Education, No.2 Shida Road, Changhua City 500, Taiwan
hhwu@cc.ncue.edu.tw

Abstract. To eliminate patients' complaints and provide better services, a case of Lutune Christian Hospital in Taiwan is illustrated. This hospital used importance-performance analysis to identify 11 out of 21 critical factors by its internal questionnaire using a Likert five-point scale from patients' viewpoints. Later, quality function deployment was applied to develop eight improvement strategies by taking into account eleven critical factors through a case management team consisting of ten psychiatric physicians, five clinical psychologists, and one case manager officially formed by this case hospital. In addition, each strategy for clinical psychology services has been expanded into a standard operating procedure to consistently meet patients' needs. The case manager plays an important role to provide a one-stop service. The satisfaction surveys are conducted during the process, and the surveys results are reviewed quarterly in the departmental meeting for improvement opportunities.

Keywords. Clinical psychology, psychiatric patient, quality function deployment, case management team, satisfaction survey, importance-performance analysis.

1. Introduction

Hargrove and Spaulding [1] stated that clinical psychology is involved heavily in understanding and treating the disorders associated with chronic mental illness. Ogunfowora and Drapeau [2] reported that many developed countries such as Canada, United Kingdom, and United States have increasing needs in clinical psychologists because these counties are experiencing growing changes to their populations due to immigration. This trend has implications for clinical psychologists to be sensitive to issues pertaining to multiculturalism and diversity in practices [2, 3]. In Taiwan, the needs for clinical psychologists have been increased due to fast-pace life style, working pressure, illness, ageing, and environment [4].

The tensed emotion tends to cause depression, anxiety, and irritability which negatively affect patients' jobs and life [4]. With the improvement of medical care in Taiwan, the focus has been changed from cure to prevention particularly in psychological health rather than physical health solely [4, 5]. In this research, a case study in Lutune Christian Hospital, one of the hospitals in Changhua Christian Medical Foundation in Changhua, Taiwan, is presented. This hospital is a psychiatric teaching hospital and intends to serve psychiatric patients and promote community mental health. In order to provide better medical services, Lutune Christian Hospital sets up a framework of clinical

psychology services to improve patients' satisfaction through patients' surveys and quality function deployment (QFD).

This study is organized as follows. Section 2 briefly reviews quality function deployment and its implementation procedures. A case study in how Lutune Christian Hospital uses patients' surveys and quality function deployment to identify critical quality items in clinical psychology services is depicted in Section 3. Besides, a case management flowchart in hospital clinical psychology services established by this hospital is addressed. Finally, conclusions are summarized in Section 4.

2. Review of Quality Function Deployment

Quality function deployment is one of the quality systems tools that can be used to relentlessly improve processes and products by fulfilling customer requirements [6, 7]. Jiang et al. [8] pointed out that QFD is the most complete, systematic, and convincing method to design products with the quality that fulfills customer requirements. The major strength of QFD is to focus and coordinate skills within an organization, from product planning and process design to manufacturing and delivering goods to customers [9]. That is, the specifications of quality requirements and deployments of quality for products and services should be started as early as possible in the life cycle [10, 11]. The most commonly seen QFD as shown in Figure 1 consists of four phases, including product planning also known as house of quality (HOQ), parts deployment, process planning, and production planning [12, 13].

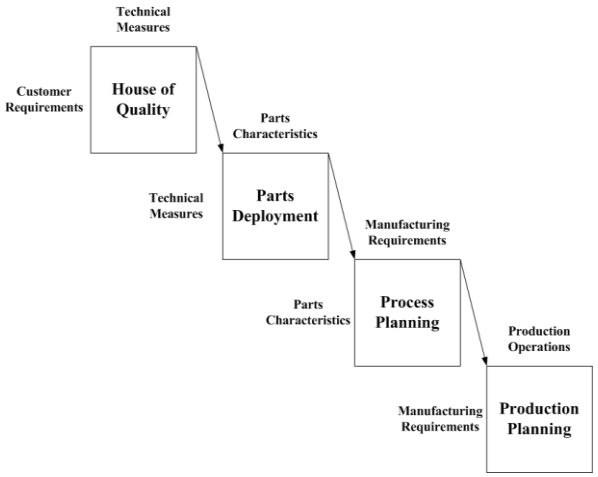


Figure 1 The Four-Phase Quality Function Deployment

In the first phase, HOQ links the voice of the customer (customer requirements) to technical measures through which detailed processes and production plans can be developed in other phases of QFD [14]. In the second phase, parts deployment is to translate important technical measures into parts characteristics, while process planning is to translate important parts characteristics into process operations in the third phase. Finally, production planning is to translate key process operations into day to day production requirements [15]. For most organizations, HOQ is the most important phase and most QFD studies only concentrate on the first phase since the structures of the other three phases are essentially the same to HOQ

[15, 16]. Besides, for the service industries, only HOQ is more practical to be used in order to improve service quality [17]. HOQ lays a foundation that products should be designed to meet customer needs such that the marketing department, design engineers, and manufacturing staffs should work closely from the time the product is first conceived [18].

The HOQ implementation (Figure 2) is composed of six steps, i.e., (1) customer needs (WHATs), (2) planning matrix, (3) technical measures (HOWs), (4) relationship matrix between WHATs and HOWs, (5) technical correlation matrix, and (6) technical matrix [19]. Chan and Wu [14] summarized that HOQ begins with customer needs and later customer competitive evaluations are performed by asking customers to evaluate the company's product compared with its major competitors' similar products in terms of the product's performance based on customer needs. In the third step, customer needs should be further transformed into technical measures. HOWs can be represented as methods, company measures, design requirements, substitute quality characteristics, and engineering characteristics which are related to and measure WHATs. The fourth step is to perform the relationship between WHATs and HOWs, which is a vital step because the final analyses depends heavily on the relationship matrix. The relationship is typically evaluated by analyzing to what extend the technical measure could technically relate to and influence the customer needs [15, 16]. Technical correlation matrix is to assess which HOWs are interrelated and how strong these relationships are through engineering analysis and experience. Finally, the purpose of technical matrix is to contain much technical information linked to both customer needs and parts characteristics, the input of the second phase in QFD by providing the initial rank order of technical measures. In addition, competitive technical assessments can be performed as well to compare the company's performance and its competitors' performance in terms of technical measures [14, 15].

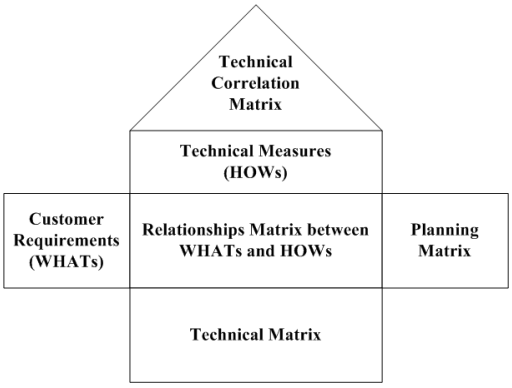


Figure 2 Six Steps of House of Quality

3. A Case Study

Before the description of this case study, a brief history of Lutune Christian Hospital is introduced. In the beginning, both neurology and psychiatry were in department of neuropsychiatry in Changhua Christian Hospital (CCH), which is the third oldest hospital (established in 1896) in Taiwan history. Later, former superintendent of CCH, Dr. David Landsborough IV, believed that neurological disorder and mental illness were different and these two kinds of diseases should be diagnosed and treated with different specialists. In 1967, CCH began to offer service for mentally ill patients and train residents of CCH as

psychiatrists. In 1972, an independent department of psychiatry was formed, which is the first independent department of psychiatry in Taiwan. Psychiatric outpatient services started in February 1972 and acute psychiatric ward began to work nine months later.

In 2005, Changhua Christian Hospital became a medical center and teaching hospital, and department of psychiatry was the largest in terms of patients, psychiatrists, psychologists, social workers, occupation therapists, and psychiatric nursing staff in central Taiwan area. Because the service could not fulfill the needs of patients, former superintendent Dr. Chiau-seng Huang encouraged the psychiatric team to organize the future plan, which was the beginning of Lutune Christian Hospital. Lutune Christian Hospital, approved by Department of Health, Executive Yuan in Taiwan in July 2004, was officially opened in October 2008 with physical and psychological medicine, psychiatry, neurology, family medicine, sleep disorders, and addiction clinics. At present, the number of employees is 150, and a total of inpatient beds are 370, including 110 beds of acute ward, 180 beds in general chronic mental illness, and 80 psychiatric day hospital beds.

The purpose of Lutune Christian Hospital is to be a psychiatric teaching hospital, which provides comprehensive, longitudinal, and holistic care to fulfill the necessity of the psychiatric patients. In the beginning, there was a lack of one-stop service center for psychiatric patients which result in poor service quality. For instance, by looking the current process of clinical psychology services, the waiting and message passing between the clinical psychology service suggested by psychiatric physicians and beginning of psychology service often results in waste of resources and complaints from patients. The complaints are composed of unavailability of the psychology service in a particular day, complaining services in outpatients' sessions only, and unchangeable treatment period in a timely basis. That is, the overall perceived service quality was relatively poor. In order to improve such circumstances, vice administrative chief decided to make changes. The framework and procedures of improving clinical psychology services are described below.

To better understand patients' needs, the questionnaire originally existed and used in this hospital was adjusted for psychiatric purposes. There are two parts of the questionnaire. The first part is demographic information, including identity, gender, age, education, living area, and types of non-drug treatment. The second part consists of the structure, process, and results of healthcare services with twenty one questions. Each patient was asked to evaluate the importance and satisfaction (performance) of every question by a Likert five-point scale, where five represents "very importance" or "very satisfactory, while one indicates "very unimportance" or "very unsatisfactory".

The questionnaire was issued by outpatient nursing staff to mental and physical outpatient directly for each psychiatrist visit. If the patients were unable to fill out the survey, their families were asked to answer those questions. From December 2011 to February 2012, thirty questionnaires have been issued but only twenty eight questionnaires are valid. The demographic information about these twenty eight questionnaires is summarized in Table 1. The detailed information including the questions and their averages in importance and performance is provided in Table 2. It is worth noting that "improved symptoms after the treatment", "patiently listening to the patients by treatment staff", "careful examination and detailed description by treatment staff", and "enough outpatients' appointment hours" are the four most important factors from patients' viewpoints, while "patiently listening to the patients by treatment staff", "comfort and quiet treatment space", and "treatment staff with professional competence" are the three most important factors in terms of satisfactory.

Table 1 The Demographic Information about Twenty Eight Questions

Information	Item	Frequency	Percentage
Identity	Patients	26	93%
	Families	2	7%
Gender	Male	17	61%
	Female	11	39%
Age	29 years old and below	8	29%
	30-39 years old	8	29%
	40-49 years old	7	25%
	50-59 years old	2	7%
	60 years old and above	3	11%
Education	Illiterate	1	4%
	Elementary school	4	14%
	Junior high school	4	14%
	Senior high school	7	25%
	College and above	12	43%
Type of non-drug treatment	Psychotherapy	5	18%
	Biofeedback	0	0%
	Parenting education	0	0%
	Psychological assessment	4	18%
	Outpatient dispensary	18	61%
	Others (Counseling)	1	4%

Table 2 The Information about Importance and Performance of Twenty One Questions

Item	Importance (Average)	Performance (Average)
1. Convenient location of hospital (close to home)	3.714	3.714
2. Convenient transportation	3.679	3.714
3. Free parking	3.893	3.929
4. Adequate and convenient parking space	3.786	3.821
5. Organized floor and clinic layout with clear directions	3.750	3.714
6. Waiting space is clean and private	3.964	3.929
7. Comfort and quiet treatment space	3.786	4.036
8. Shorter interval for the next scheduled treatment	3.857	3.571
9. Rapid and efficient treatment process	3.964	3.714
10. Good reputation of treatment staff	3.929	3.857
11. Careful examination and detailed description by treatment staff	4.071	4.000
12. Treatment staff with professional competence	3.964	4.036
13. Patiently listening to the patients by treatment staff	4.107	4.071
14. Improved symptoms after the treatment	4.214	3.786
15. Enough outpatients' appointment hours	4.071	3.714
16. A variety of treatment staff to be chosen	3.857	3.643
17. Guided personnel with appointment in the hospital	3.714	3.714
18. Clear description of the leaflets provided with appointment	3.750	3.857
19. Reminding service for appointment	3.571	3.536
20. Quick and proper response for treatment enquire	4.000	3.857
21. Good communication with patients before and after treatment	4.000	3.750
Grand Average	3.888	3.808

By applying importance-performance analysis, the decision maker can easily track those more critical factors for continuous improvement, particularly with items in Quadrant IV (concentrate here) followed by Quadrant III (low priority) [20, 21]. Specifically, items falling in Quadrant IV have low performance but high importance, which can be viewed as major weaknesses. Immediate attention for improvement is required. On the other hand, items located in Quadrant III have both low performance and importance, which can be classified as minor weaknesses for improvement [21]. From Figure 3, items 1, 7, 8, and 17 fall in Quadrant III, and items 6, 9, 11, 12, 14, 20, and 21 are in Quadrant IV. Therefore, these eleven items should be improved in order to provider better healthcare services.

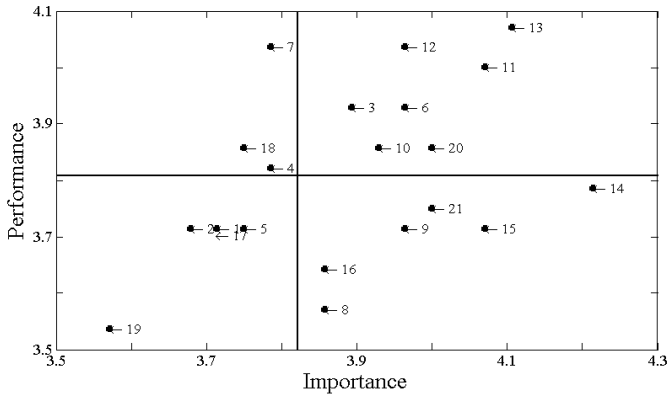


Figure 3 Importance-Performance Analysis Results

The purpose of identifying these critical factors is to not only improve those items but also establish standard operating procedures (SOP) to relentlessly provide excellence services to patients. Thus, vice administrative chief decided to set up a management care team in clinical psychological cases. Table 3 provides specific information about team members. Ten out of sixteen members are psychiatric physicians, and five members are clinical psychologists. Both psychiatric physicians and clinical psychologists have different professional but they have something in common, i.e., psychological treatment. That is, psychiatric physicians use interviews for the clinical diagnosis of psychological treatment, and clinical psychologists use interviews for psychological assessment and psychological treatment. It is worth to note that a case manager was hired to be in charge of one-stop services during the framework setup. In fact, he is "the bridge" between the hospital and patients when services are requested by patients.

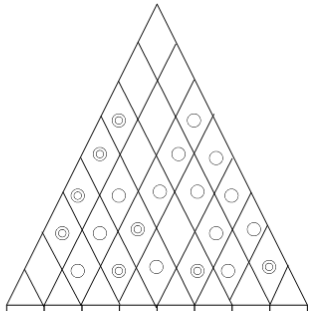
Table 3 The Case Management Team

	psychiatric physician	clinical psychologist	case manager
Number of members	10	5	1
Average practice years	18.9	15.3	6.8
Professional qualification	Yes	Yes	No
Psychotherapy experience	Yes	Yes	No
Parenting education experience	Yes	Yes	No
Biofeedback experience	No	Yes	No
Psychological assessment experience	No	Yes	No

When the team is formed and critical factors are identified, quality function deployment can be used to improve service. Because "convenient location of hospital (close to home)" identified as a minor weakness cannot be improved easily, this factor is removed in this stage. Thus, in Step 1, ten instead of eleven critical factors are the input. The weights of these critical factors are their respective importance values depicted in Table 2. In Step 2, in order to generate appropriate technical measure items to reflect patients' needs, this team regularly meets once per week from the last week of March 2012 to the last week of April 2012. By adopting multiple assessments and analyses for those patients' needs and reviewing the current clinical psychology services, this team with brainstorming has concluded eight major improvement strategies, including "the establishment of clinical psychological case management office", "the establishment of standard operating procedures for case management", "unified regulation of treatment period", "unified regulation of treatment space", "unified arrangement of the treatment", "unified assist in arranging treatment mode", "regularly review of satisfaction survey results for improvement opportunity", and "regularly review of quality indicators of SOP".

Step 3 is to evaluate the relationship between patients' needs and technical measures. This team uses consensus to determine the relationships with the notations of ◎, ○, and △ representing strong, medium, and weak relationships, respectively. The blank between patients' needs and technical measures indicates there is no relationship between them. To further quantify the relationships, the numerical figures of 9, 3, and 1 represent the respective strong, medium, and weak relationships. In the roof as discussed in Step 4, if any pair of technical measures have strong and positive relationship, the notation of ◎ is applied. If the relationship is positive, the notation of ○ is shown. In contrast, if the relationship is negative, the notation of × is used. Finally, if the relationship is strong and negative, the notation of ∗ is used. The major purpose is to examine if any pair of technical measures are a paradox. In this case, there is no any paradox existed in any pair of technical measures.

In the planning matrix as described in Step 5, the hospital intends to meet patients' needs at this stage. Thus, there is no need to make any comparison among different hospitals in terms of patients' requirements. Finally, Step 6 is to compute the overall importance of eight technical measures and then prioritize technical measures when the resources are limited. The detailed clinical psychological care of house of quality is shown in Figure 4.



Patients' Requirements	Importance	The establishment of clinical psychological case management office	The establishment of standard operating procedures for case management	Unified regulation of treatment period	Unified regulation of treatment space	Unified arrangement of the treatment	Unified assist in arranging treatment mode	Regularly review of satisfaction survey results for improvement opportunity	Regularly review of quality indicators of SOP
6. Waiting space is clean and private	3.964	9	9	9	9	9	3	9	9
7. Comfort and quiet treatment space	3.786	9	3	3	3			9	9
8. Shorter interval for the next scheduled treatment	3.857	9	9	9	9	9	9	9	9
9. Rapid and efficient treatment process	3.964	9	9	9	9	9	9	9	9
11. Careful examination and detailed description by treatment staff	4.071					3	3	3	3
12. Treatment staff with professional competence	3.964					3	3	3	3
14. Improved symptoms after the treatment	4.214					3	3	3	3
17. Guided personnel with appointment in the hospital	3.714	9	9	9	9	9	3	9	9
20. Quick and proper response for treatment enquire	4.000	3	3	3	3	9	3	3	3
21. Good communication with patients before and after treatment	4.000	3	3	3	3	9	3	3	3
Total Score		197.565	174.849	174.849	174.849	248.238	154.170	234.312	234.312
Priority of Technical Measures		4	5	5	5	1	8	2	2

Figure 4 Clinical Psychological Care of House of Quality

From Figure 4, "unified arrangement of the treatment" is the most important technical measure, followed by "regularly review of satisfaction survey results for improvement opportunity", and "regularly review of quality indicators of SOP". In order to improve service relentlessly, regularly review of satisfaction survey results for improvement opportunity and regularly review of quality indicators of SOP should be placed in second highest priority. Besides, unified arrangement of the treatment, unified regulation of treatment period and unified regulation of treatment space can actively improve waiting time and provide customized services. Thus, these three technical measures can be integrated to be part of case manager's duties. The standard operating procedures of this case management are depicted in Figure 5. With the establishment of clinical psychological case management

office along with the integration of case manager's duties, one-stop clinical psychology services can be provided to further eliminate the complaints from patients.

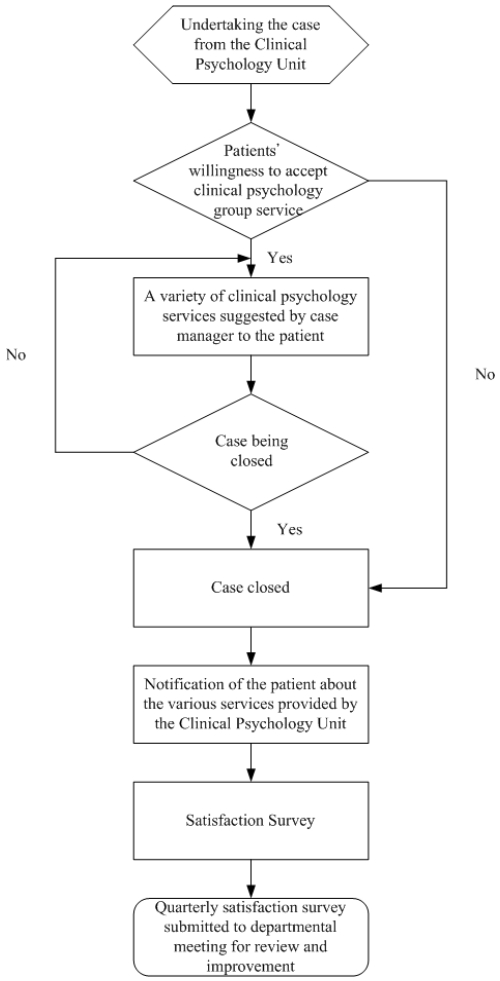


Figure 5 The Flowchart of Case Manager's Duties

4. Conclusions

A case study is presented by depicting how Lutune Christian Hospital first uses importance-performance analysis to identify critical factors from patients' surveys, and then forms a case management team to apply quality function deployment to generate eight improvement strategies. In order to provide better, consistent, and customized services to patients, each improvement strategy has been expanded into SOP such that the clinical psychology services can meet a wide variety of patients' needs. Most importantly, the case manager plays an important role for each case and monitors the progress of clinical psychology services. The satisfactory surveys are conducted during the process and will be reviewed quarterly in the departmental meeting for improvement opportunities.

References

1. Hargrove D.S., W.D. Spaulding, Training psychologists for work with the chronically mentally ill, *Community Mental Health Journal*, 24(4), 283-295, 1988.
2. Ogunfowora B., M. Drapeau, Comparing counseling and clinical psychology practitioners: similarities and differences on theoretical orientations revisited, *International Journal for the Advancement of Counselling*, 30(2), 93-103, 2008.
3. Williams P.E., G. Turpin, G. Hardy, Clinical psychology service provision and ethnic diversity within the UK: A review of the literature, *Clinical Psychology and Psychotherapy*, 13(6), 324-338, 2006.
4. Chen H.J., Y.R. Chang, The perspective of past, present and future of clinical health psychology in Taiwan, *Research in Applied Psychology*, 40, 99-119, 2008.
5. Liu Y.C., Y.L. Lan, S.P. Kao, The situated learning approach in clinical psychology education: the use of challenge of the standardized patient training method, *Formosa Journal of Mental Health*, 24(2), 247-278, 2011.
6. Kuo T.C., H.H. Wu, Green products development by applying grey relational analysis and green quality function deployment, *International Journal of Fuzzy Systems*, 5(4), 229-238, 2003.
7. Liu C.H., H.H. Wu, A fuzzy group decision-making approach in quality function deployment, *Quality and Quantity*, 42(4), 527-540, 2008.
8. Jiang J.C., M.L. Shiu, M.H. Tu, QFD's evolution in Japan and the West, *Quality Progress*, 40(7), 30-37, 2007.
9. Raharjo H., M. Xie, A.C. Brombacher, A systematic methodology to deal with the dynamics of customer needs in Quality Function Deployment, *Expert Systems with Applications*, 38(4), 3653-3662, 2011.
10. Hauser J.R., D. Clausing, The house of quality, *Harvard Business Review*, 66(2), 63-73, 1988.
11. Wu H.H., J.I Shieh, Applying a Markov chain model in quality function deployment, *Quality and Quantity*, 42(5), 665-678, 2008.
12. Wu H.H., A.Y.H. Liao, P.C. Wang, Using grey theory in quality function deployment to analyse dynamic customer requirements, *International Journal of Advanced Manufacturing Technology*, 25(11&12), 1241-1247, 2005.
13. Kuo T.C., H.H. Wu, J.I Shieh, Integration of environmental considerations in quality function deployment by using fuzzy logic, *Expert Systems with Applications*, 36(3P2), 7148-7156, 2009.
14. Chan L.K., M.L. Wu, Quality function deployment: a comprehensive review of its concepts and methods, *Quality Engineering*, 15(1), 23-35, 2002-03.
15. Chan L.K., M.L. Wu, A systematic approach to quality function deployment with a full illustrative example, *Omega*, 33(2), 119-139, 2005.
16. Wu H.H., J.I Shieh, Applying repertory grids technique for knowledge elicitation in quality function deployment, *Quality and Quantity*, 44(6), 1139-1149, 2010.
17. Chang H., D. Kim, A quality function deployment framework for the service quality of health information websites, *Healthcare Informatics Research*, 16(1), 6-14, 2010.
18. Tan K.C., X.X. Shen, Integrating Kano's model in the planning matrix of quality function deployment, *Total Quality Management*, 11(8), 1141-1151, 2000.
19. Wu H.H., J.I Shieh, Using a Markov chain model in quality function deployment to analyse customer requirements, *International Journal of Advanced Manufacturing Technology*, 30(1-2), 141-146, 2006.
20. Hu H.-Y., Y.-C. Lee, T.-M. Yen, C.-H. Tsai, Using BPNN and DEMATEL to modify importance–performance analysis model – A study of the computer industry, *Expert Systems with Applications*, 36(6), 9969-9979, 2009.
21. Wu H.H., Y.T. Tang, J.W. Shyu, A case of applying importance-performance analysis in identifying key success factors to develop marketing strategies, *Quality and Quantity*, 44(6), 1207-1218, 2010.

Evaluating Patient Safety Culture from Viewpoints of Physicians and Registered Nurses – A Case of a Regional Hospital in Taiwan

Yii-Ching Lee[1], Chin-Lun Hsu[2], Hsin-Hung Wu[3], Liang-Po Hsieh[4], Shao-Jen Weng[5], Chih-Hsuan Huang[6]

[1]Department of Medical Quality Management, Cheng Ching General Hospital, Taichung City, Taiwan
yiiching.lee@gmail.com

[2]Department of Business Administration, National Changhua University of Education, Changhua, Taiwan
scottray3@hotmail.com

[3]Department of Business Administration, National Changhua University of Education, Changhua, Taiwan
hhwu@cc.ncue.edu.tw

[4]Department of Medical Quality Management, Cheng Ching General Hospital, Taichung City, Taiwan
lphsieh624@yahoo.com.tw

[5]Department of Industrial Engineering and Enterprise Information, Tunghai University, Tacihung City, Taiwan
sjweng@thu.edu.tw

[6]Department of Industrial Engineering and Enterprise Information, Tunghai University, Taichung City, Taiwan
tititacer@gmail.com

Abstract. The purpose of this study is to observe the attitudes from medical physicians and registered nurses toward patient safety culture. This study uses t-test and one-way analysis of variance to evaluate how different demographic variables including gender, age, job status, job position, experience in organization, and experience in position impact nine dimensions of patient safety culture. Bonferroni test is performed for post hoc analysis except for gender and job position with only two levels. The results are summarized below. Teamwork climate is more important for males than females. Age has significant impacts on eight dimensions. Five dimensions were influenced significantly by experience in organization. Finally, experience in position affects teamwork across hospital units and hospital handoffs and transitions.

Keywords. Patient safety culture, analysis of variance, Bonferroni test, demographic variable, Chinese version of hospital survey.

1. Introduction

Patient safety culture plays an important role for hospitals to relentlessly improve patient safety [1]. Schutz et al. [2], Shie et al. [3] and Zuniga et al. [4] stated that better attitude toward patient safety could result in positive influences on shorter stay, fewer prescription errors, less ventilator-associated pneumonia, fewer blood-stream and urinary tract infections, and lower mortality. Wagner et al. [5] also stated that hospitals with a more open culture and reflective attitude toward errors and patient safety could reduce the number

of accidents and failures. Therefore, it is critically important for healthcare organizations to regularly measure employees' safety attitudes since a positive patient safety culture indicates that healthcare organizations place patient safety as one of the highest priorities [3, 6, 7].

Patient safety can be measured by safety culture surveys [8]. The hospital survey on patient safety culture (HSPSC), originally developed by Sexton et al. [9], with good validity and reliability plays an important role to assess opinions of patient safety culture from healthcare organizations' staffs [1, 7]. In hospitals, the core staffs are medical physicians and registered nurses. Thus, it would be of interest to observe the attitudes from medical physicians and registered nurses toward patient safety culture. Because medical physicians and registered nurses have different demographic information such as gender, age, job status, job position, experience in organization, and experience in position, it also would be worth to analyze how different demographic information impacts patient safety culture.

2. Definition of Corporate Sustainability

The HSPSC has six dimensions and each dimension is defined as below [9]. Teamwork climate is defined as perceived quality of collaboration between personnel. Safety climate is the perceptions of a strong and proactive organizational commitment to safety. Job satisfaction is defined as the positivity of the work experience. Stress recognition measures how performance is impacted by stressors. Perception of management is defined as the approval of managerial action. Finally, working conditions is from the perceived quality of the work environment and logistical support such as staffing and equipment. Respondents who are health organizations' staffs are required to answer questions by rating each question based on a five-point Likert scale ranging from strongly agree to strongly disagree or frequency such as never, rarely, sometimes, most of the time, and always [5, 9].

Taiwan Joint Commission on Hospital Accreditation developed the Chinese version of patient safety culture survey in 2008 based on the short form of the HSPSC in 2006 by using forward and backward translation to check the quality of the translation and by pilot-testing and discussion by an expert panel for intelligibility and applicability of the items [5, 10]. Originally with six dimensions and thirty questions, Taiwan Joint Commission on Hospital Accreditation further incorporated three hospital-level aspects of safety culture from Agency for Healthcare Research and Quality. That is, three items, four items, and four items are from hospital management support for patient safety, teamwork across hospital units, and hospital handoffs and transitions, respectively [11]. Therefore, the Chinese version of HSPSC consists of nine dimensions and 41 questions. Table 1 summarizes the information of nine dimensions and forty one questions.

Table 1 The Chinese Version of Hospital Survey on Patient Safety Culture

Teamwork Climate	
1	Nurse input is well received in this clinical area.
2	In this clinical area, it is difficult to speak up if I perceive a problem with patient care.
3	Disagreements in this clinical area are resolved appropriately (i.e., not who is right, but what is best for the patient).
4	I have the support I need from other personnel to care for patients.
5	It is easy for personnel here to ask questions when there is something that they do not understand.
6	The physicians and nurses here work together as a well-coordinated team.
Safety Climate	
7	I would feel safe being treated here as a patient.

8	Medical errors are handled appropriately in this clinical area.
9	I know the proper channels to direct questions regarding patient safety in this clinical area.
10	I receive appropriate feedback about my performance.
11	In this clinical area, it is difficult to discuss errors.
12	I am encouraged by my colleagues to report any patient safety concerns I may have.
13	The culture in this clinical area makes it easy to learn from the errors of others.
Job Satisfaction	
14	I like my job.
15	Working here is like being part of a large family.
16	This is a good place to work.
17	I am proud to work in this clinical area.
18	Morale in this clinical area is high.
Stress Recognition	
19	When my workload becomes excessive, my performance is impaired.
20	I am less effective at work when fatigued.
21	I am more likely to make errors in tense or hostile situations.
22	Fatigue impairs my performance during emergency situations (e.g. emergency resuscitation, seizure).
Perception of Management	
23	Management supports my daily efforts.
24	Management doesn't knowingly compromise patient safety.
25	I get adequate, timely info about events that might affect my work.
26	The levels of staffing in this clinical area are sufficient to handle the number of patients.
Working Conditions	
27	Problem personnel are dealt with constructively by our unit.
28	This hospital does a good job of training new personnel.
29	All the necessary information for diagnostic and therapeutic decisions is routinely available to me.
30	Trainees in my discipline are adequately supervised.
Hospital Management Support for Patient Safety	
31	Management is doing a good job.
38	The actions of hospital management show that patient safety is a top priority.
39	Hospital management seems interested in patient safety only after an adverse event happens.
Teamwork across Hospital Units	
32	Hospital units do not coordinate well with each other.
34	There is good cooperation among hospital units that need to work together.
36	It is often unpleasant to work with staff from other hospital units.
40	Hospital units work well together to provide the best care for patients.
Hospital Handoffs and Transitions	
33	Things "fall between the cracks" when transferring patients from one unit to another.
35	Important patient care information is often lost during shift changes.
37	Problems often occur in the exchange of information across hospital units.
41	Shift changes are problematic for patients in this hospital.

3. Research Method

Under the current system, the Chinese version of HSPSC with nine dimensions and forty one questions is conducted in a yearly basis and the original data set should be submitted electronically to Taiwan Joint Commission on Hospital Accreditation through the website. By removing incomplete questionnaires, the number of effective questionnaires

from physicians and nurses in 2010 is 361. The demographic information of physicians and nurses including gender, age, job status, job position, experience in organization, and experience in position is shown in Table 2.

Table 2 Demographic Information of Physicians and Nurses

Demographic Variable	Frequency	Percentage
Job Position		
Physician	36	10.0
Nurse	325	90.0
Gender		
Male	41	11.4
Female	320	88.6
Age		
Less than 20 years old	0	0
21-30 years old	163	45.2
31-40 years old	143	39.6
41-50 years old	50	13.9
51-60 years old	5	1.4
61 years old and above	0	0
Job Status		
Full-time	323	89.5
Part-time	25	6.9
Agency	8	2.2
Contract	5	1.4
Experience in Organization		
Less than 6 months	30	8.3
6-11 months	38	10.5
1-2 years	55	15.2
3-4 years	59	16.3
5-10 years	85	23.5
11-20 years	89	24.7
21 years and above	5	1.4
Experience in Position		
Less than 6 months	37	10.2
6-11 months	43	11.9
1-2 years	62	17.2
3-4 years	68	18.8
5-10 years	94	26.0
11-20 years	57	15.8
21 years and above	0	0

In order to analyze how different demographic variables influence patient safety culture from an overall viewpoint, independent sample t-test for mean difference and one-way analysis of variance (ANOVA) with $\alpha = 0.05$ are applied to nine dimensions of patient safety culture. The scale for each question ranges from one to five. In this study, the total score for each dimension will be used for analyses. For instance, there are six questions in teamwork climate. Thus, the total score of teamwork climate is from six to thirty. By the same token, the total scores of the other dimensions can be computed. Since the core hospital staffs in this case hospital are medical physicians and registered nurses, this study intends to observe the attitudes from medical physicians and registered nurses toward patient safety

culture. If the p-value is less than $\alpha = 0.05$, Bonferroni test is performed for post hoc analysis except for gender and job position with only two levels.

4. Results

The independent sample t-test for mean difference of gender with the assumption of the same variance is provided in Table 3. The p-value of teamwork climate is 0.030, which is less than 0.05, indicating that different genders have different perceptions on teamwork climate. The mean values of males and females are 22.85 and 21.28, respectively. As a result, males have better perceptions on teamwork climate than females.

Table 3 t-Test for Mean Difference of Gender on Nine Dimensions

Dimension	Mean (Male)	Mean (Female)	t-value	p-value
Teamwork Climate	22.85	21.28	2.185	.030
Safety Climate	25.10	24.30	1.018	.310
Job Satisfaction	17.95	16.90	1.529	.127
Stress Recognition	15.63	14.53	1.861	.064
Perception of Management	13.24	12.98	.554	.580
Working Conditions	14.05	13.34	1.476	.141
Hospital Management Support for Patient Safety	9.80	9.88	-.224	.823
Teamwork across Hospital Units	12.78	12.44	.837	.403
Hospital Handoffs and Transitions	11.34	11.59	-.604	.546

The ANOVA results on age are provided in Table 4, where eight out of nine dimensions except for hospital handoffs and transitions are significant. The Bonferroni test on these eight dimensions is depicted in Table 5. Due to limited space, we only highlight the significant mean differences. No age differences on stress recognition. However, age of 21-30 years old has lower satisfaction than age of 41-50 years old in teamwork climate, safety climate, job satisfaction, working conditions, hospital management support for patient safety, and teamwork across hospital units. Moreover, age of 31-40 years old has lower satisfaction than age of 41-50 years old in safety climate, job satisfaction, perception of management, working conditions, hospital management support for patient safety, and teamwork across hospital units. Generally, although age of 51-60 years old has the highest perceptions than age of 41-50 years old, the sample of size for age of 51-60 years old is five. By excluding the smaller sample of size of age of 51-60 years old, age of 41-50 years old has the highest satisfaction on these nine dimensions even though some value differences are not statistically significant with $\alpha = 0.05$. In other words, the older the physicians and nurses are in this case hospital, the better perceptions they have in patient safety culture.

Table 4 Age on Nine Dimensions

Dimension		Sum of Squares	df	Mean Square	F	Sig.
Teamwork Climate	Between Groups	260.127	3	86.709	4.660	.003
	Within Groups	6643.369	357	18.609		
	Total	6903.496	360			
Safety Climate	Between Groups	348.470	3	116.157	5.393	.001
	Within Groups	7689.458	357	21.539		
	Total	8037.928	360			
Job Satisfaction	Between	368.583	3	122.861	7.510	.000

	Groups Within Groups Total	5840.281 6208.864	357 360	16.359		
Stress Recognition	Between Groups Within Groups Total	118.659 4535.058 4653.717	3 357 360	39.553 12.703	3.114	.026
Perception of Management	Between Groups Within Groups Total	83.108 2998.881 3081.989	3 357 360	27.703 8.400	3.298	.021
Working Conditions	Between Groups Within Groups Total	109.316 2910.684 3020.000	3 357 360	36.439 8.153	4.469	.004
Hospital Management Support for Patient Safety	Between Groups Within Groups Total	45.342 1355.539 1400.881	3 357 360	15.114 3.797	3.980	.008
Teamwork across Hospital Units	Between Groups Within Groups Total	92.182 2025.951 2118.133	3 357 360	30.727 5.675	5.415	.001
Hospital Handoffs and Transitions	Between Groups Within Groups Total	30.886 2245.834 2276.720	3 357 360	10.295 6.291	1.637	.181

Table 5 Post Hoc Analysis on Age

Dimension	I (Age)	J (Age)	Mean Difference (I-J)	Sig.
Teamwork Climate	21-30	41-50	-1.946	.033
Safety Climate	21-30	41-50	-2.458	.007
	31-40	41-50	-2.447	.009
Job Satisfaction	21-30	41-50	-2.723	.000
	31-40	41-50	-2.315	.003
Perception of Management	31-40	41-50	-1.359	.027
Working Conditions	21-30	41-50	-1.439	.012
	31-40	41-50	-1.506	.009
Hospital Management Support for Patient Safety	21-30	41-50	-.941	.018
	31-40	41-50	-1.063	.006
Teamwork across Hospital Units	21-30	41-50	-1.412	.002
	31-40	41-50	-1.378	.003

Table 6 summarizes that no job status differences on nine dimensions. Job position has significant differences in teamwork climate, stress recognition, and teamwork across hospital units as shown in Table 7. In fact, physicians perceive better than nurses statistically.

Table 6 Job Status on Nine Dimensions

Dimension		Sum of Squares	df	Mean Square	F	Sig.
Teamwork Climate	Between Groups	69.859	3	23.286	1.217	.304
	Within Groups	6833.637	357	19.142		
	Total	6903.496	360			
Safety Climate	Between Groups	75.832	3	25.277	1.133	.335
	Within Groups	7962.096	357	22.303		
	Total	8037.928	360			
Job Satisfaction	Between Groups	41.559	3	13.853	.802	.493
	Within Groups	6167.305	357	17.275		
	Total	6208.864	360			

Stress Recognition	Between Groups	9.783	3	3.261	.251	.861
	Within Groups	4643.934	357	13.008		
	Total	4653.717	360			
Perception of Management	Between Groups	9.492	3	3.164	.368	.776
	Within Groups	3072.497	357	8.606		
	Total	3081.989	360			
Working Conditions	Between Groups	15.053	3	5.018	.596	.618
	Within Groups	3004.947	357	8.417		
	Total	3020.000	360			
Hospital Management Support for Patient Safety	Between Groups	5.541	3	1.847	.473	.702
	Within Groups	1395.340	357	3.909		
	Total	1400.881	360			
Teamwork across Hospital Units	Between Groups	5.332	3	1.774	.300	.826
	Within Groups	2112.811	357	5.918		
	Total	2118.133	360			
Hospital Handoffs and Transitions	Between Groups	1.281	3	.427	.067	.977
	Within Groups	2277.439	357	6.374		
	Total	2276.720	360			

Table 7 t-Test for Mean Difference of Job Position on Nine Dimensions

Dimension	Physician	Nurse	t-value	p-value
Teamwork Climate	23.28	21.25	2.655	.008
Safety Climate	25.25	24.30	1.151	.251
Job Satisfaction	18.28	16.88	1.923	.055
Stress Recognition	16.14	14.49	2.634	.009
Perception of Management	13.28	12.98	.588	.557
Working Conditions	13.92	13.37	1.082	.280
Hospital Management Support for Patient Safety	9.94	9.86	.239	.811
Teamwork across Hospital Units	13.25	12.40	2.011	.045
Hospital Handoffs and Transitions	11.67	11.55	.255	.799

Table 8 indicates that experience in organization has statistically significant differences in safety climate, job satisfaction, perception of management, teamwork across hospital units, and hospital handoffs and transitions. Post hoc analysis summarized in Table 9 shows that the staffs with 11-20 years of experience in organization perceive better safety climate than the staffs with 1-2 and 3-4 years. For job satisfaction, staffs with less than 6 months and with 11-20 years of experience in organization have better job satisfaction than staffs with 1-2 years. Staffs with 11-20 years of experience in organization have better perception of management than staffs with 6-11 months and 1-2 years. In addition, staffs with 11-20 years have better teamwork across hospital units than staffs with 1-2 years. It is worth noting that staffs with less than 6 months have better teamwork across hospital units than staffs with 1-2, 3-4, and 5-10 years. Finally, for hospital handoffs and transitions, staffs with less than 6 months feel more satisfied than staffs with 3-4 and 5-10 years.

Table 8 Experience in Organization on Nine Dimensions

Dimension		Sum of Squares	df	Mean Square	F	Sig.
Teamwork Climate	Between Groups	226.657	6	37.776	2.003	.065
	Within Groups	6676.839	354	18.861		
	Total	6903.496	360			

Safety Climate	Between Groups	419.782	6	69.964	3.251	.004
	Within Groups	7618.146	354	21.520		
	Total	8037.928	360			
Job Satisfaction	Between Groups	366.001	6	61.000	3.696	.001
	Within Groups	5842.864	354	16.505		
	Total	6208.864	360			
Stress Recognition	Between Groups	129.446	6	21.574	1.688	.123
	Within Groups	4524.271	354	12.780		
	Total	4653.717	360			
Perception of Management	Between Groups	160.442	6	26.740	3.240	.004
	Within Groups	2921.547	354	8.253		
	Total	3081.989	360			
Working Conditions	Between Groups	94.047	6	15.674	1.896	.081
	Within Groups	2925.953	354	8.265		
	Total	3020.000	360			
Hospital Management Support for Patient Safety	Between Groups	46.288	6	7.715	2.016	.063
	Within Groups	1354.593	354	3.827		
	Total	1400.881	360			
Teamwork across Hospital Units	Between Groups	137.057	6	22.843	4.082	.001
	Within Groups	1981.076	354	5.596		
	Total	2118.133	360			
Hospital Handoffs and Transitions	Between Groups	87.110	6	14.518	2.347	.031
	Within Groups	2189.610	354	6.185		
	Total	2276.720	360			

Table 9 Post Hoc Analysis on Experience in Organization

Dimension	I	J	Mean Difference (I-J)	Sig.
Safety Climate	1-2 years	11-20 years	-2.859	.008
	3-4 years	11-20 years	-2.531	.027
Job Satisfaction	1-2 years	Less than 6 months	-3.097	.018
	1-2 years	11-20 years	-2.332	.019
Perception of Management	6-11 months	11-20 years	-1.838	.022
	1-2 years	11-20 years	-1.713	.012
Teamwork across Hospital Units	1-2 years	Less than 6 months	-2.155	.002
	3-4 years	Less than 6 months	-1.934	.006
	5-10 years	Less than 6 months	-1.582	.037
	1-2 years	11-20 years	-1.243	.049
Hospital Handoffs and Transitions	3-4 years	Less than 6 months	-2.000	.008
	5-10 years	Less than 6 months	-1.647	.041

Experience in position has significant impacts on two out of nine dimensions, namely teamwork across hospital units and hospital handoffs and transitions as shown in Table 10. Table 11 summarizes post hoc analysis for teamwork across hospital units and hospital handoffs and transitions. Staffs with less than 6 months in experience in position feel more satisfied with teamwork across hospital units than staffs with 1-2, 3-4, and 5-10 years. By the same token, staffs with less than 6 months have more positive impacts on hospital handoffs and transitions than staffs with 3-4 and 5-10 years. An interesting finding is that staffs with less than 6 months in experience in position have the highest satisfaction values in both teamwork across hospital units and hospital handoffs and transitions even though some mean differences are not statistically significant. In contrast, staffs with 11-20 years have lowest satisfaction in both teamwork across hospital units and hospital handoffs and transitions.

Table 10 Experience in Position on Nine Dimensions

Dimension		Sum of Squares	df	Mean Square	F	Sig.
Teamwork Climate	Between Groups	68.145	5	13.629	.708	.618
	Within Groups	6835.351	355	19.255		
	Total	6903.496	360			
Safety Climate	Between Groups	152.056	5	30.411	1.369	.235
	Within Groups	7885.872	355	22.214		
	Total	8037.928	360			
Job Satisfaction	Between Groups	179.937	5	35.987	2.119	.063
	Within Groups	6028.927	355	16.983		
	Total	6208.864	360			
Stress Recognition	Between Groups	77.664	5	15.533	1.205	.306
	Within Groups	4576.053	355	12.890		
	Total	4653.717	360			
Perception of Management	Between Groups	75.217	5	15.043	1.776	.117
	Within Groups	3006.772	355	8.470		
	Total	3081.989	360			
Working Conditions	Between Groups	38.178	5	7.636	.909	.475
	Within Groups	2981.822	355	8.399		
	Total	3020.000	360			
Hospital Management Support for Patient Safety	Between Groups	29.357	5	5.871	1.520	.183
	Within Groups	1371.524	355	3.863		
	Total	1400.881	360			
Teamwork across Hospital Units	Between Groups	94.770	5	18.954	3.326	.006
	Within Groups	2023.363	355	5.700		
	Total	2118.133	360			
Hospital Handoffs and Transitions	Between Groups	80.291	3	16.058	2.595	.025
	Within Groups	2196.429	355	6.187		
	Total	2276.720	360			

Table 11 Post Hoc Analysis on Experience in Position

Dimension	I	J	Mean Difference (I-J)	Sig.
Teamwork across Hospital Units	Less than 6 months	1-2 years	1.741	.008
	Less than 6 months	3-4 years	1.669	.010
	Less than 6 months	5-10 years	1.448	.029
Hospital Handoffs and Transitions	Less than 6 months	3-4 years	1.536	.040
	Less than 6 months	5-10 years	1.608	.014

5. Conclusions

This study evaluates how different demographic variables impact patient safety culture from the viewpoints of physicians and nurses. The results show that teamwork climate is more important for males. Age has statistically significant impacts on eight dimensions except for hospital handoffs and transitions. Specifically, age of 41-50 years old has the second best perception on eight dimensions by post hoc analysis. Safety climate, job satisfaction, perception of management, teamwork across hospital units, and hospital handoffs and transitions were influenced significantly by experience in organization. Staffs with 11-20 years in experience in organization are more satisfied in safety climate, job satisfaction, and teamwork across hospital units. Staffs with less than 6 months have higher

perception in job satisfaction, teamwork across hospital units, and hospital handoffs and transitions. Moreover, experience in position influences teamwork across hospital units and hospital handoffs and transitions. Staffs with less than 6 months in experience in position are more satisfied.

Acknowledgement

This study was partially supported by Cheng Ching General Hospital – Chung Kang Branch, Taichung City, Taiwan with grant number of CH10200159.

References

1. Chen, I.C., H.H. Li, Measuring patient safety culture in Taiwan using the hospital survey on patient safety culture (HSOPSC), *BMC Health Services Research*, 10, 152, 2010.
2. Schutz, A.L., M.A. Counte, S. Meurer, Development of a patient safety culture measurement tool for ambulatory health care settings: analysis of content validity, *Health Care Management Science*, 10(2), 139-149, 2007.
3. Shie, H.G., W.C. Lee, H.F. Hsiao, H.L. Lin, L.L. Yang, F. Jung, Patient safety attitudes among respiratory therapists in Taiwan, *Respiratory Care*, 56, 1924-1929, 2011.
4. Zuniga, F., D. Schwappach, S. De Geest, R. Schwendimann, Psychometric properties of the Swiss version of the nursing home survey on patient safety culture, *Safety Science*, 55, 88-118, 2013.
5. Wagner, C., M. Smits, J. Sorra, C.C. Huang, Assessing patient safety culture in hospitals across countries, *International Journal for Quality in Health Care*, 25(3), 213-221, 2013.
6. Wilson, D.S., Registered nurses' collective safety organising behaviours: the association with perceptions of patient safety culture, *Journal of Research in Nursing*, 18(4), 320-333, 2012.
7. Fujita, S., K. Seto, S. Ito, Y. Wu, C.C. Huang, T. Hasegawa, The characteristics of patient safety culture in Japan, Taiwan and the United States, *BMC Health Services Research*, 13, 20, 2013.
8. Deilkas, E., D. Hofoss, Patient safety culture lives in departments and wards: Multilevel partitioning of variance in patient safety culture, *BMC Health Services Research*, 10, 85, 2010.
9. Sexton, J.B., R.L. Helmreich, T.B. Neilands, K. Rowan, K. Vella, J. Boyden, P.R. Roberts, E.J. Thomas, The safety attitudes questionnaire: Psychometric properties, benchmarking data, and emerging research, *BMC Health Services Research*, 6, 44, 2006.
10. Lee, W.C., H.Y. Wung, H.H. Liao, C.M. Lo, F.L. Chang, P.C. Wang, A. Fan, H.H. Chen, H.C. Yang, S.M. Hou, Hospital safety culture in Taiwan: a nationwide survey using Chinese version safety attitude questionnaire, *BMC Health Services Research*, 10, 234, 2010.
11. Lee, L.J., C.H. Hsu, J.C. Hong, Explore the impact of hospital restructuring medical practitioners of patient safety culture attitudes, *Journal of Health Management*, 10(2), 189-199, 2012.

Effect of a Regional-Industrial Promotion Policy through Open Innovation - Analysis of Effects of Open Source Software in Japanese IT Companies' Business Growth, with Special Reference to the "Region"

Tetsuo Noda

Shimane University, 1060 Matsue, Japan
nodat@soc.shimane-u.ac.jp

Abstract. Traditional regional-industrial promotion policies in Japan have mainly been based on the cluster theory presented by Porter (1998). On the other hand, owing to the rapid development of information technologies, networks linking business enterprises with each other and outside resources have progressed significantly. Even the field of research and development is now open to the world and connected with external resources, a concept similar to the "Open Innovation" theory put forward by Chesbrough (2009). In response, Matsue City, a typical peripheral region in Japan, has proposed a regional industry promotion project since 2006 built around the open source programming language Ruby, not by forming a regional industry cluster but by organizing a cooperation that extends beyond regional borders. In this paper, we regard this regional IT industry promotion policy by Matsue City, as a process of advancing Open Innovation in local area. We analyze how OSS effects Japanese IT companies' business growth both through simple use and by deeper engagement as a stakeholder in OSS community, with special reference to the "Region". Result of the questionnaire survey to IT companies in Japan, it has become clear that industrial promotion policy by local government advances the practical use of OSS and contribution to OSS communities. At the same time, it is supposed that a regional-industry policy has not affected the indicators of business growth yet.

Keywords. Industrial Promotion Policy, Open Innovation, Open Source, OSS, Regionality

1. Introduction

Traditional regional-industrial promotion policies in Japan have mainly been based on the cluster theory presented by Porter (1998). On the other hand, owing to the rapid development of information technologies, networks linking business enterprises with each other and outside resources have progressed significantly. Even the field of research and development is now open to the world and connected with external resources, a concept similar to the "Open Innovation" theory put forward by Chesbrough (2009). In response, Matsue City, a typical peripheral region in Japan, with a population of less than 200,000 people, has proposed a regional-industrial promotion project since 2006 built around the open source programming language Ruby, not by forming a regional industry cluster but by organizing a cooperation that extends beyond regional borders.

As a consequence of this regional-industrial promotion policy, sales in the IT industry in this city increased by 57.9% from 2006 to 2010, compared to a negative 3.7% growth rate in the rest of Japan during the same period. Furthermore, the number of employees in this

city increased by 77.8% from 2006 to 2010, while the national average was 12.9% over the same period (The Ministry of Economy Trade and Industry of Japan, 2008-2012).

In this paper, we regard this regional-industrial promotion policy by Matsue City, as a process of advancing Open Innovation in local area. Chesbrough (2003) describes "Open Innovation" which blurs boundaries between business enterprises, and by joining internal resources and external resources together, extra economic values for all parties concerned is generated. This development style is essentially the same as the longer established OSS development style. OSS is developed by a "Community" of stakeholders, which may be structured as a Bazaar style or a Cathedral style, it tends to be open for all developers, software engineers and business enterprises with an interest in participation, and they can participate or withdraw any stage in the overarching process. From the perspective of businesses engaging as stakeholders in this field, they join a community beyond the separated confines of their own organization in order to absorb the fruit of innovation from third parties, who participate for similar reasons in turn. It is essentially a situation of enlightened self-interest. One immediate consideration from this perspective is that the development of OSS technology inherently reduces costs for each stakeholder, with the complete burden of development being shared by all contributors. Conceptualizing OSS technologies as platforms on which products or services can be delivered, it is easy to extrapolate that OSS contributions therefore can be directly tied into reductions in the cost of bringing new products and services to market, and therefore provides a market benefit through what can be called a leverage effect.

Kunai (2010) categorizes open source business model into the "Three-step Model" from the involvement of business enterprises. As moving up the ladder, though the cost of the development increases, business enterprises can increase the economic effects as shown in **Fig. 1**.

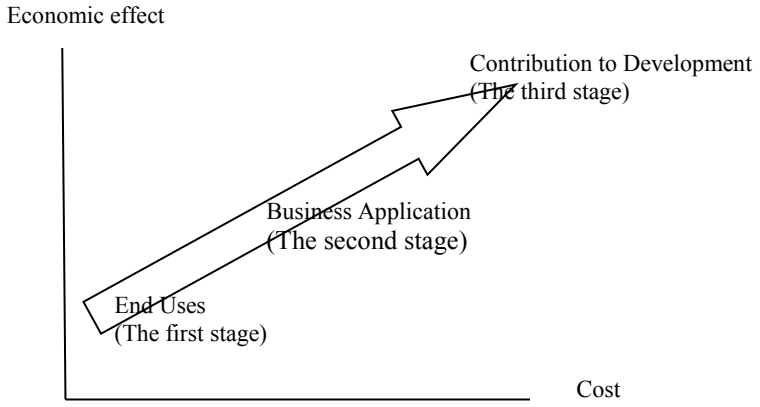


Fig. 1: Three-step Business Model of Open Source

In the first stage, business enterprises use OSS as End Users, and they only use OSS in the same was as proprietary software. Their primary purpose is cost reduction, but economic effect is very low. In the second stage they use OSS in a more engaged manner, expanding functional features they need, constructing application software, serving support for their customers, and integrating systems. In this stage, the economic effect is comparatively higher than that of the first stage, though cost rises because of the demand of manpower and equipment to launch and sustain these derivative businesses. In the third stage they participate in the "mainstream" development process of OSS, and bring forth the highest economic effect. They contribute to the "Community" by providing physical support and financial backing. The development style of this stage is different from to stage two, primarily because they develop software in association with other companies, including their competitors. Business enterprises become to be able to reduce the cost of the demand of manpower and equipment. Moreover, developing with OSS engineers and other companies, they are able to acquire the "Leverage Effect". Thus the underlying hypothesis is that process of "Open Innovation" enables business enterprises to absorb the fruits of the "Community" of OSS. It is assumed that background factor of the success of the regional-industrial promotion policy by Matsue City depends on the process of "Open Innovation" beyond boundaries of enterprises and regions.

This paper analyzes how OSS effects Japanese IT companies' business growth both through simple use and by deeper engagement as a stakeholder in OSS community, with special reference to the "Region".

2. Study Methodology and Questionnaire Survey

The methodology in this study is to investigate the effect on the business growth by OSS utilization and contribution in Japanese IT companies. As is described by Kunai, we assume, "The more IT companies contribute to OSS communities, the more they are able to acquire economic effect". Now, this paper tries to establish this hypothesis by the questionnaire survey to IT companies in Japan.

During 2012, questionnaire survey slips were sent to 642 companies which accede to Information Industry Association in Japan, and 191 companies gave us replies (collection rate: 29.8%). Among respondent companies, the number of companies of Big Cities (Tokyo, Osaka, and Fukuoka) is 90, that of Local Cities (except Matsue) is 78, and that of Matsue City is 28.

The survey was conducted in the form of a questionnaire containing the items shown in box 1. In the survey we questioned the utilization and contribution of low-level OSS (such as Linux, Database technologies, Programming Languages, etc.). Application-level software (such as ERP, CMS, CRM, etc.) is excluded, because case examples of development of such software are rare in Japanese IT companies. All questions are selected from among alternatives, discrete data.

Box 1: OSS Questionnaire Survey Slips towards Japanese IT Companies

Company profile:
- Q1. Home City
- Q2. Inauguration of Business
- Q3. Main Business Service
- Q4. Capital Stock
- Q5. Number of Employee
- Q6. Number of Developers (programmers, software engineers, etc.)
- Q7. Sales Amount
- Q8. Growth Rate of Sales (present period)
- Q9. Prospect of Sales Growth Rate (subsequent period)
- Q10. Growth of Employee Number (present period)
- Q11. Prospect of Employee Number's Growth Rate (subsequent period)

Utilization of OSS: (rate of utilization)
- Q12. Utilization of Linux
- Q13. Utilization of Apache HTTP Server
- Q14. Utilization of Database technologies (MySQL, PostgreSQL, etc.)
- Q15. Utilization of Programming Language Ruby
- Q16. Utilization of Other Programming Languages (Perl, Python, PHP, etc.)
- Q17. Utilization of Ruby on Rails

Contribution to OSS Communities:
(amount of direct investments and manpower costs of OSS engineers inside company)
- Q18. Contribution to Linux
- Q19. Contribution to Apache HTTP Server
- Q20. Contribution to Database technologies (MySQL, PostgreSQL, etc.)
- Q21. Contribution to Programming Language Ruby
- Q22. Contribution to Other Programming Languages (Perl, Python, PHP, etc.)
- Q23. Contribution to Ruby on Rails

3. Regional Comparison of Utilization of OSS

Most Japanese IT companies use OSS in their business field, especially the Linux operating system core components and various Database technologies (MySQL, PostgreSQL, etc.). Better part of Japanese IT companies uses these OSS more than quarter part of their system components (**Fig. 2**).

At the same time, the use rate of Japanese-origin technology like Ruby and its American-based development framework Ruby on Rails are unexpectedly low. It is because that in the questionnaire survey the poll of IT companies contains wide ranges, including system integrators, software developers, and network service companies. Ruby use is currently limited within the field of web applications development, along with Ruby on Rails.

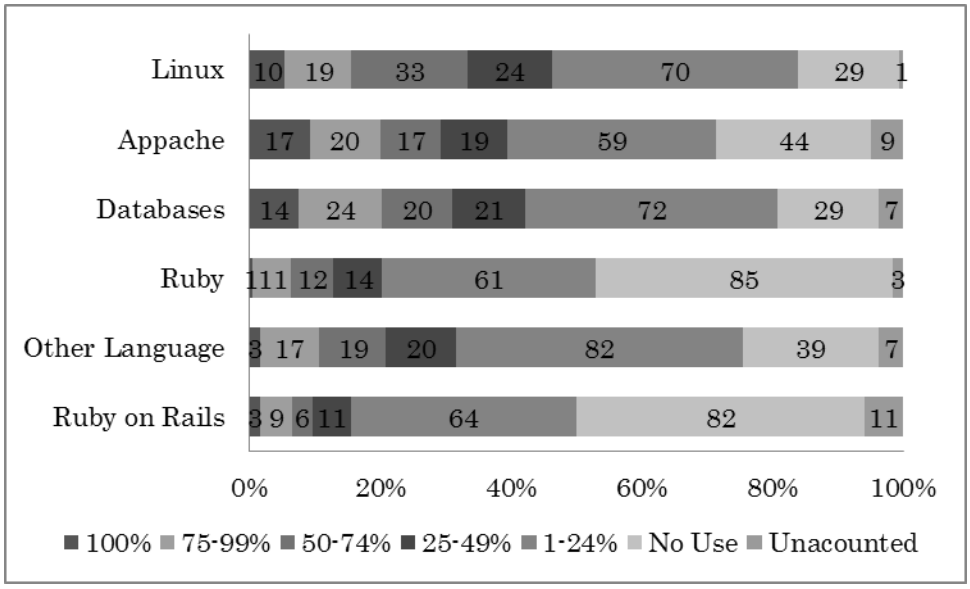


Fig. 2: Utilization of OSS in Japanese IT Companies n=191

Next, we draw a comparison of utilization of OSS among Big Cities (Tokyo, Osaka, and Fukuoka), Local Cities (except Matsue) and Matue City, taking Linux and Ruby as examples. In Matsue City, in spite of local peripheral area in Japan, the percentage of companies using Linux is high on a par with Big Cities (**Fig. 3**). This tendency is also proved in other OSS. Moreover, the percentage of companies using Ruby is extremely-high (**Fig. 4**).

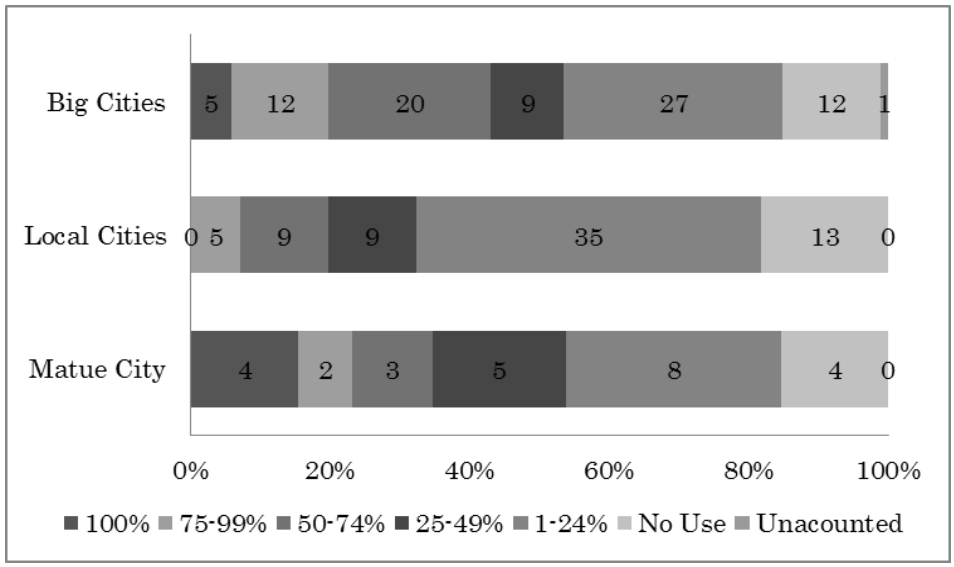


Fig. 3: Comparison of utilization of Linux by Region

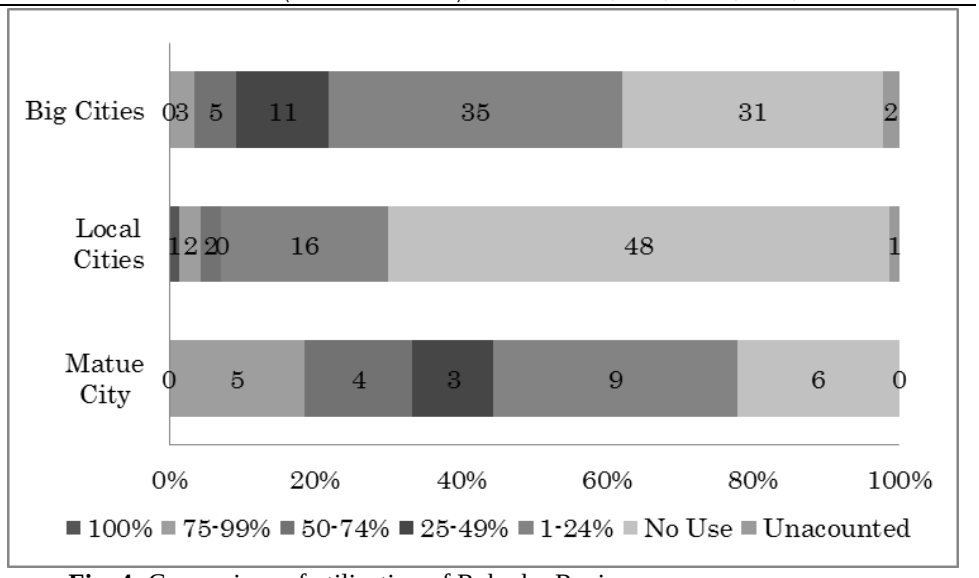


Fig. 4: Comparison of utilization of Ruby by Region

4. Regional Comparison of Contribution to OSS

The survey also revealed that the percentage of companies which contribute to OSS communities is relatively low in Japan (**Fig. 5**). The fact that most of Japanese IT companies use OSS without contributing to OSS development process might show that they are positioned as "free riders". However, the survey also confirmed that, on average, about 20% of IT companies contribute to OSS development process.

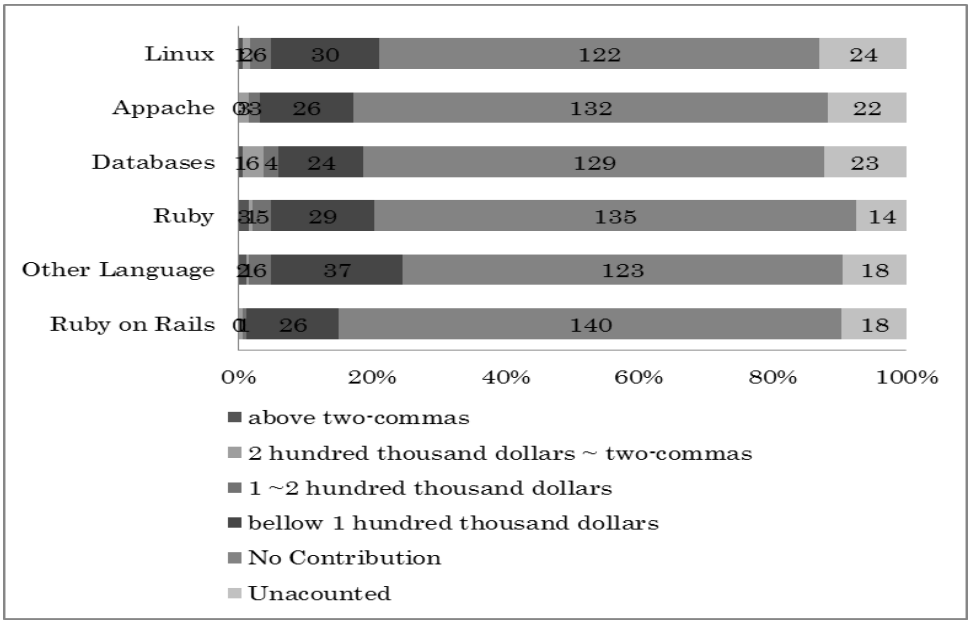


Fig.5: Contribution to OSS communities in Japanese IT Companies n=191

We also draw a comparison of contribute to OSS communities taking Linux and Ruby as examples. By the same token, in Matsue City the percentage of companies contributing to Linux is high on a par with Big Cities (**Fig. 6**), and the percentage of companies contributing to Ruby is extremely-high (**Fig. 7**).

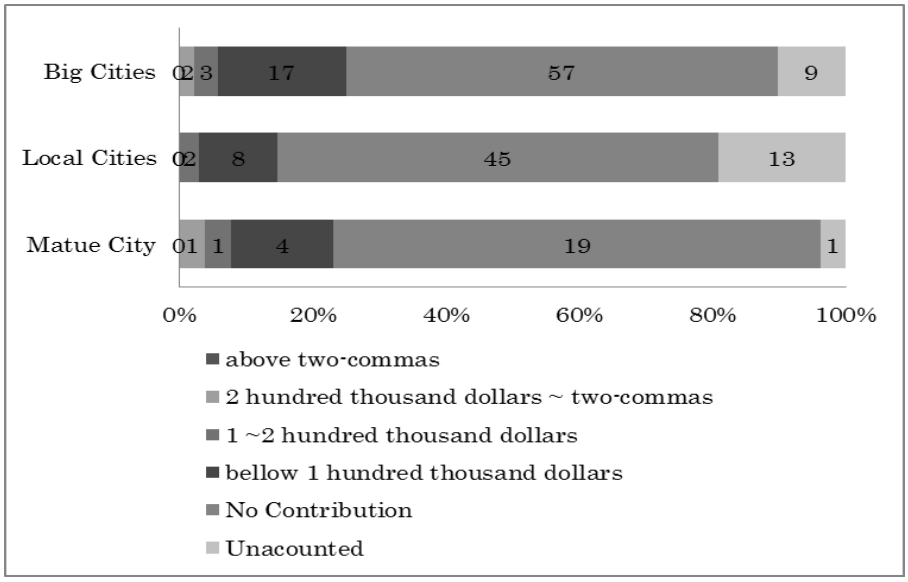


Fig.6: Comparison of contribution to OSS communities of Linux by Region

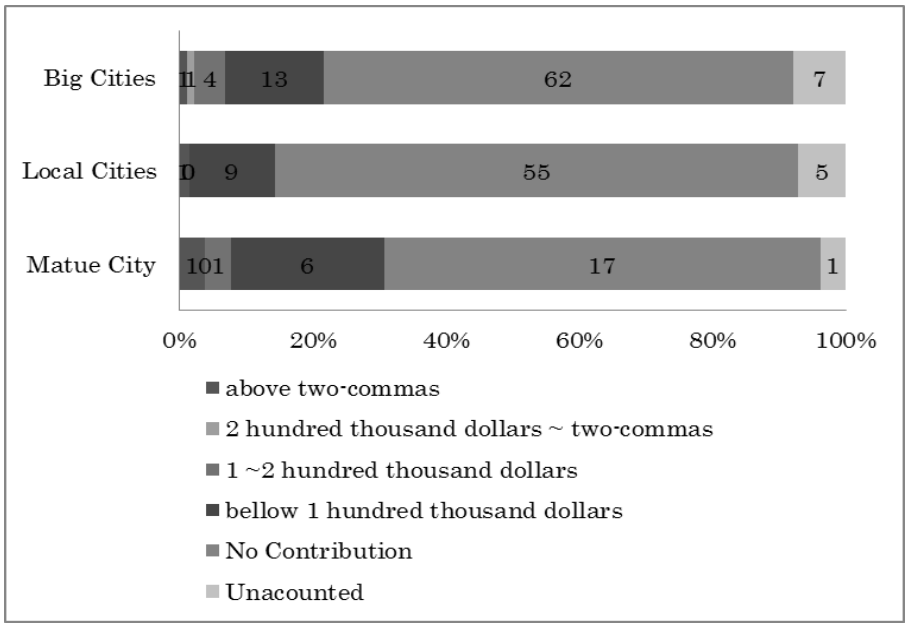


Fig.7: Comparison of contribution to OSS communities of Ruby by Region

As just described, IT companies in Matsue City not only use OSS but also contribute to OSS communities, especially Ruby. A regional-industrial promotion project of Matsue City has increased the amount of sales and the number of employees in the IT industry in this city as discussed previously. Now, we analyse whether this fruits of the regional-industrial promotion project comes from the utilization and contribution of OSS, or more specifically Regional Open Innovation. "The more regional IT companies contribute to OSS communities, the more they are able to acquire economic effect"

5. Effect on Business Growth by Utilization of OSS and Contribution of OSS

In order to consider the effect on business growth by the IT companies' utilization and contribution of OSS, and to consider the regionality, we do logit analysis letting indexes of corporate management as dependent variables. Four items of corporate management are each simplified whether it demonstrates an upward trend (1), or not (0). Utilization of OSS and contribution to OSS communities are put as dependant dummy variables. Big City Dummy (Big City (1), Others (0)) and Matsue Dummy (Matsue (1), Others (0)) are also added as control variables. **Table 1** shows the result of logit analysis.

[Independent Variables (1,0)]	[Dependent Variables]
Growth Rate of Sales	Utilization of OSS (6 items)
- Present Period, Subsequent Period	Contribution to OSS (6 items)
Growth Rate of Employee Number	Regions (2 items, Big Cities and Matsue City)
- Present Period, Subsequent Period	

Linux Utilization affects positively Growth rate of Sales and Employee Number in Present Period. Growth rate of Employee Number in Present Period is affected by Apache, positively by contribution, but negatively by utilization.

As for Regions, Big City Dummy positively affects Growth of Employee Number in Subsequent Period (+ 10% level of significance), but Matsue Dummy is negative (5% level of significance). Other items indicate no significance. It is supposed that the fruits of the regional-industrial promotion project by Matsue City have not come from the utilization and contribution of OSS at present.

Table 1: Logit Analysis of Growth Rate of Sales and Growth of Employee Number

		Growth Rate of Sales			Growth of Employee Number		
		Present Period	Subsequent Period (Prospect)		Present Period	Subsequent Period (Prospect)	
OSS Utilization	Linux	.437 +	.417 +		.815 **	.269	
	Apache	-.175	-.185		-.665 *	-.163	
	Databases	.068	.056		-.110	-.085	
	Ruby	.083	.168		.375	.304	
	Other Languages	.046	.249		.159	.182	
	Ruby on Rails	-.009	-.071		-.205	-.092	
OSS Contribution	Linux	-.154	.298		-.109	-.048	
	Apache	.989	.002		1.456	.190	
	Databases	-.179	.355		-.588	.554	
	Ruby	-.319	-.677		.001	-.012	
	Other Languages	-.407	-.361		-.710	-1.063	
	Ruby on Rails	.089	.507		-.120	.153	
Regions	Big Cities Dummy Big City(1) Others(0)	.649	.587		.540	1.042	+
	Matsue Dummy Big City(1) Others(0)	.101	.515		-.208	-.278	*
constant numbers		-1.730	-2.171		-1.921	-1.872	
Wald Statistics		11.570 **	15.134		23.093	23.093	
Log Likelihood		158.144	150.120		139.511	144.541	
Cox-Shell R2		.109	.136		.159	.127	
Nagelkerke R2		.150	.189		.227	.182	

** 1% level of significance, * 5% level of significance, + 10% level of significance

6. Conclusion

It has become commonplace for business enterprises to use OSS in their business. The logic we understand as framing this such engagement is that the competitive edge that comes from technical advantages delivered by using OSS, and - using the same logic - it is therefore indispensable for them to contribute or participate in the development process of OSS as Kunai proposes. Our data shows that major OSS, like Linux, Apache, MySQL, and PostgreSQL, are still utilization objects for Japanese IT companies. They have been able to get a competitive edge only by the utilization of OSS, and contribution to OSS projects or communities has not been linked to the business growth for them. At the same time, The industrial promotion policy by local government (Matsue City) advances not only the practical use of OSS but also the contribution to OSS communities. However, it is supposed that the regional industry policy has not affected the indicators of business growth yet.

We excluded the survey of application-level software (such as ERP, CMS, CRM, etc.). For the future, case examples of development of such software are expected to increase in Japanese IT companies. Moreover, to survey the effect on business growth we take on growth rate of sales and growth of employee number as indicators of business growth. There are also other indicators to estimate business growth. These are our research challenges for the future.

References

1. Chesbrough, H., Open Innovation: The New Imperative for Creating And Profiting from Technology, Harvard Business School Press. 2003.
2. Chesbrough, H., Open Business Models: How To Thrive In The New Innovation Landscape, Harvard Business School Press. 2006.
3. Chesbrough, H., Wim, V., and West, J. , Open Innovation: Researching a New Paradigm, Oxford University Press. 2008.
4. Kunai, T., Three-stage approach of Linux Development, Let's go by Linux vol.2. 2010.
5. The Ministry of Economy Trade and Industry of Japan, Specific Service Industry Investigation of Actual Conditions, 2008-2012.
6. Piore, M.J. and Sabel, C.F., The Second Industrial Divide, Basic Books, 1984.
7. Porter, M.E., On Competition, Harvard Business School Press, Boston, 1998.

A General Framework to Coordinate a Decentralized Supply Network

Atour Taghipour

Faculty of International Business, University of Le Havre,
atour.taghipour@univ-lehavre.fr

Abstract. Supply chain is a complex system composed of two or more independent companies. Management of supply chain is generally focuses on the coordination of these independent companies under asymmetric information. The complexity increased by adding new companies to the supply chain. The question is to how coordination can be achieved in a supply chain with more than two partners when independent entities do not want exchange their strategic information to other members of this network. This research addresses the problem of coordination in a supply chain with more than two partners and proposes a general framework to discover the space of coordination. Using incentive systems to influence their partner's planning, this approach requires only a minimum level of information sharing. Our objective is to develop and to demonstrate that this approach can produce near optimal solution, compared to centralized-planning.

Keywords. Supply chain management, coordination, network, operations planning.

1. Introduction

Supply chains are complex networks which includes numerous independent organizations connected through upstream and downstream linkages under form of information, physical and financial flows. These organizations have different objectives and operations planning domains. A key issue in Supply Chain Management is the coordination of decision making systems of these independent organizations.

The literature of supply chain coordination can be divided into two main streams of research: centralized vs. decentralized approaches. A centralized approach can theoretically produces optimal to all partners. However, its implementation requires a high degree of information exchange. At the other side, in decentralized coordination partners are considered as separate economic entities that make their own planning decisions independently.

This paper addresses the problem of supply chain network coordination using a minimum level of information sharing. In this paper, we extend an approach first introduced in Taghipour and Frayret (2012), which coordinates two partners of a supply chain, to consider simple form of a supply chain including one supplier and several manufacturers.

Existing frameworks of supply network coordination including a single supplier and multiple buyers focus on contract-based coordination under centralized mechanisms or cooperative structures while aspect of mutual coordination of operations planning has not studied clearly. For example, in Cachon (2003) using buy-back contracts the supplier encourages the buyers to establish optimal policies which is not possible except by providing full information necessary for globally optimization. In other work (Chen et al., 2001), a price-based coordination is proposed. In fact, to coordinate the supply chain, the supplier plays the role of a central planner and use a discount mechanism to induce the buyers to

exchange the full information. As another example, in Karabatı and Sayın (2008), to obtain complete cost information of buyers, the supplier exchanges his own cost parameters with buyers leading to vertical information sharing.

2. Problem Overview

In this work, a supplier and several manufacturers have to plan their own lot sizes, in terms of what to produce, in what quantity and when over a planning horizon. To coordinate their plans, unlike traditional centralized system, partners interact directly with each other to mutually adjust their lot-sizes in order to improve their individual and collective economic performance.

In this context, the supplier uses an incentive mechanism to incite the manufacturers to adjust their original *order plans* (OP) which is defined as the manufacturers' orders to the supplier for all products and all time periods of the planning horizon. When one partner makes a proposal, the other partner must evaluate the proposal in order to know its feasibility, as well as its impact on profit. The positive difference between the supplier's optimal proposal and the manufacturer's original OP is referred to as the *Additional Supply Plan (ASP)*, which represents how the supplier wants the manufacturer to increase its order for specific products at specific time periods.

Next, the supplier calculates the *Maximum Discount* (MD) that can be offered to the manufacturer if he accepts the supplier's optimal OP. The *Maximum Discount* is defined as the gap between the profit generated from delivering its optimal OP and the profit generated from delivering the manufacturer's original OP. Finally, using the ASP and the MD, the supplier defines and offers a *Discount Plan* (DP) to the manufacturer, which consists in offering part of the *MD* for an adjustment of the original OP equal to part of the *ASP*. In other words, if the manufacturer accepts to increase its order for specific products at specific time periods up to at least the specific portion of the ASP, than a fixed discount is offered to the manufacturer (Figure 1).

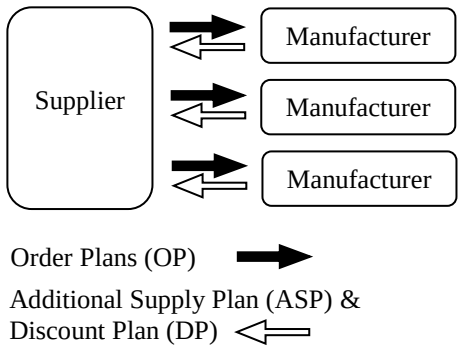


Figure 1: Exchanged information.

Unlike the pioneer approaches which require partners to exchange perfect information, the proposed approach propose to use a dynamic discount structure that is progressively adjusted in order to find a compromise OP without sharing strategic information.

3. Coordination Framework

The planning starts with the upstream planning scheme. Each manufacturer determines his locally optimal plan and transmits associated order plan to the supplier who plans based on the received order plans form all manufacturers. Resulting is the upstream planning solution. Next, the supplier generates a counter-proposal which includes an Additional Supply Plan (ASP) and a Discount Plan (DP). The associated proposals are then transmitted to the manufacturers for evaluation. Each manufacturer analyses the received supply proposal and replies by eventual counter-proposal to the supplier's proposal. The supplier analyses the impact of these counter-proposals on his local planning.

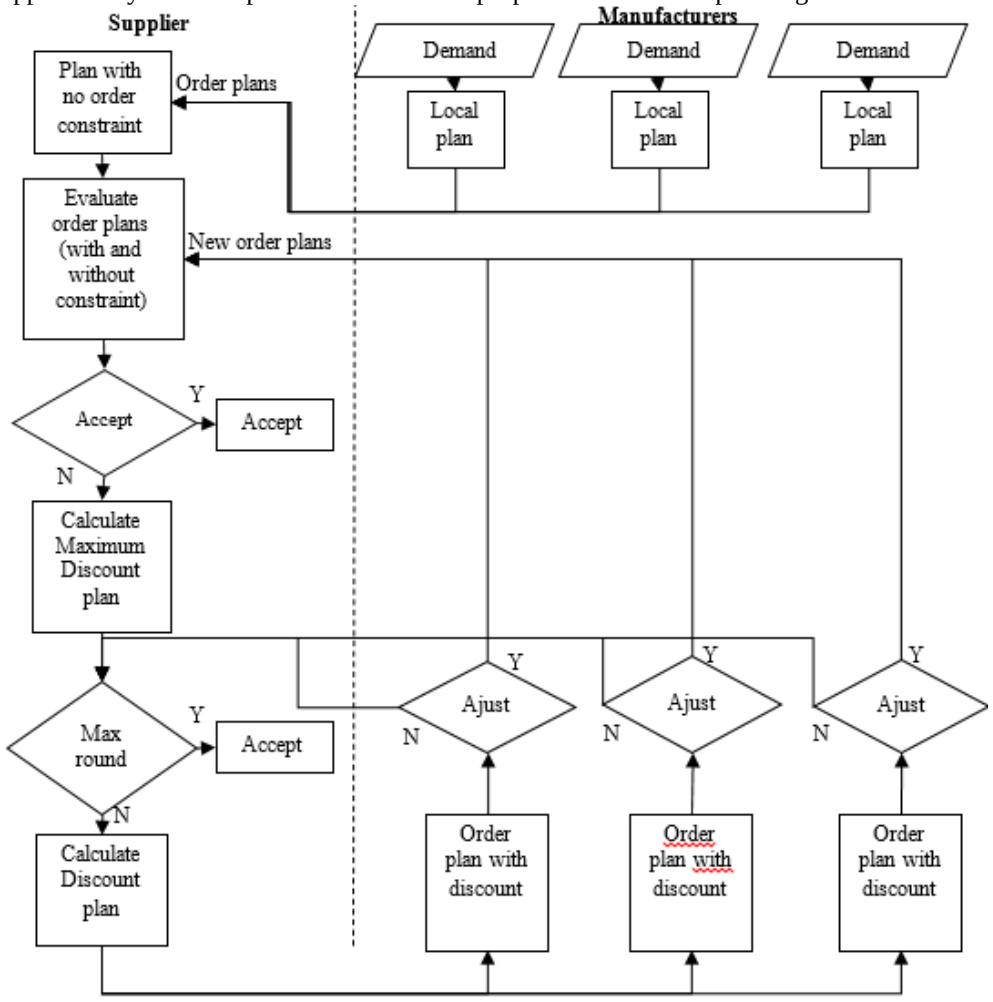


Figure 2: Coordination Framework

4. Solution Examination

Our approach is examined by considering a simple supply chain including one supplier and three manufacturers. The aim is to plan the production of two products over two planning periods. In this example, the upstream planning approach, as an initial constrained based solution, produces a total supply chain profit of 3500 (Table 1). However, as suggested in the same table, it is more profitable for the supplier to switch the production of 30 units of product 1 from period 1 to period 2, with a local potential profit increase of 2500 - 500 = 2000. This increase in supplier's profit is the result of a relaxed lot-size plan without considering the result of upstream planning.

Table 1: Constrained and relaxed order plans.

		Manufacturer 1 (Profit = 1000)		**Manufacturer 2** (Profit = 1000)		**Manufacturer 3** (Profit = 1000)		**Upstream plan** **Supplier** (Profit = 500)		**Relaxed plan** **Supplier** (Potential profit = 2500)	
		Periods		**Periods**		**Periods**		**Periods**		**Periods**	
		1	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**
Products	**1**	20	0	20	0	20	0	60	0	30	30
	2	10	0	10	0	10	0	30	0	30	0

In order to offer a discount, it is necessary to create a discount plan based on the maximum discount plan. So, the supplier computes the Additional Supply Plan and Maximum Discount Plan, which are presented in Table 2. These two matrix will be used to create incentive proposals to the each of the three manufactures.

Table 2: Supplier's Additional Supply Plan and Maximum Discount Plan.

		Additional Supply Plan		**Maximum Discount Plan**	
		Periods		**Periods**	
		1	**2**	**1**	**2**
Products	**1**	0	30	0	2000
	2	0	0	0	0

In the next stage, the supplier proposes in a first round of negotiation, 30% of this maximum discount plan ($\alpha 1 = 0.1, \alpha 2 = 0.1, \alpha 3 = 0.1$) to the three manufacturers. Each manufacturer can benefit from10% of maximum discount plan, if he increases his order plan for product 1, period 2, at least from 0 to 2 units (10% of the ASP) ($\beta = 0.1$). See Table 3.

Table **Error! No text of specified style in document.**: First round of negotiation.

		Proposed discount $\alpha 1 = 0.1$; $\alpha 2 = 0.1$; $\alpha 3 = 0.1$ & $\beta = 0.1$)		Manufacturer 1 Plan (profit = 1000)		Manufacturer 2 Plan (profit = 1000)		Manufacturer 3 Plan (profit = 1000)		Supplier Plan (profit = 500)	
		Periods		Periods		Periods		Periods		Periods	
		1	2	1	2	1	2	1	2	1	2
Products	1	0	600	20	0	20	0	20	0	20	0
	2	0	0	10	0	10	0	10	0	10	0

In this example, because the incentive is not sufficient to change the manufacturers' order plans, the supplier increases the offered discount to 60% of the Maximum Discount Plan ($\alpha 1 = 0.2, \alpha 2 = 0.2, \alpha 3 = 0.2$), for a similar increase of its order plan for product 1, period 2, from 0 to, at least, 2 units (10% of the ASP,$\beta = 0.1$), which results in three partners benefiting from improved profit compared to the original solution, the supplier and two manufacturers, as shown in Table 4.

Table 4: Second round of negotiation.

		Proposed discount $\alpha 1 = 0.2$; $\alpha 2 = 0.2$; $\alpha 3 = 0.2$ & $\beta = 0.1$)		Manufacturer 1 Plan (profit = 1000)		Manufacturer 2 Plan (profit = 1000)		Manufacturer 3 Plan (profit = 1000)		Supplier Plan (profit = 2000)	
		Periods		Periods		Periods		Periods		Periods	
		1	2	1	2	1	2	1	2	1	2
Products	1	0	1200	20	0	15	5	10	10	45	15
	2	0	0	10	0	10	0	10	0	10	0

5. Conclusion

This paper presents a general framework to coordinate a decentralized supply chain network. A particular relation between one supplier and three manufacturers is studied. This approach is based on an iterative negotiation between these partners. It is a non-hierarchical mechanism, which gives the same decision authority to both partners without any exchange of sensitive information like cost structures (Dudek and Stadtler, 2007). Using an incentive mechanism one partner encourage the manufacturers to participate in coordination process

and to address simultaneously material and financial flows. More computational results are necessary to show promising results.

References

1. Cachon, G.P., Supply chain coordination with contracts. In Handbooks in Operations Research and Management Science – Supply Chain Management, edited by S. Graves and T. de Kok, pp. 229–339, 2003 (Amsterdam: North Holland).
2. Chen, F., Federgruen, A. and Zheng, Y.-S., Coordination mechanism for a distribution system with one supplier and multiple retailers. Manage. Sci., vol. 47, 2001, pp. 693–708.
3. Dudek, G., Stadtler, H., Negotiation-based collaborative planning in divergent two-tier supply chains, Intrenational Journal of Production Research, vol. 45, n°. 2, 15 January 2007, pp. 465-484.
4. Karabati, S., Sayin, S., Single-supplier/multiple-buyer supply chain coordination: Incorporating buyers' expectations under vertical information sharing, European Journal of Operational Research, vol. 187, 2008, pp. 746–764.
5. Taghipour, A., Frayret, J.-M., Mutual adjustment search with incentive for supply chain planning coordination, International Journal of Computer Integrated Manufacturing, vol. 25, n° 10, October 2012, pp. 946-962.

Opportunities for Business Intelligence in a Cloud

Kamelia Stefanova, Dorina Kabakchieva, Stanimira Yordanova

University of National and World Economy, 1700 Sofia, Bulgaria
kstefanova@unwe.bg, dkabakchieva@unwe.bg, smira.yordanova@gmail.com

Abstract. The paper is devoted to the main aspects of Cloud Business Intelligence. Presented are the current trends of the collaboration of the two technologies – Cloud Computing and Business Intelligence. The focus is put on the specific characteristics of Business Intelligence and Software as a Service, Infrastructure as a Service, and Platform as a Service. Some of the critical challenges for Cloud Business Intelligence are discussed.

Keywords. Cloud Computing, Business Intelligence, Software as a Service, Infrastructure as a Service, Platform as a Service.

1. Introduction

The Business environment dynamic development today, more than ever puts a very high pressure on the enterprises to be profitable and competitive. Companies realize the need to introduce advanced analytical tools in order to get the right insight of their business and opportunities and be able to take the required decisions and necessary actions. Business Intelligence (BI) introduces not only new technologies and processes to deeply understand the information about the environment for strategic purposes but it is becoming a new philosophy of gaining knowledge about the business development and figuring out the right trends.

Companies from all kind of sizes consider the challenges of incorporating adequate BI solutions in their information processing systems. The theory and practice of BI is still searching for principals in helping the companies to overcome the difficulties in selecting the right approach in implementing the business intelligence in all the companies' management levels.

An increasing trend in the BI development is appearing to become the Cloud solutions. Processing huge amounts of data requires respected resources and space. Those resources are getting critical for the companies to fulfill their daily electronic operations, on one hand and on the other, to adequately process, analyse and understand the information needed to support the management process.

Cloud business intelligent solutions provide a new approach to the organization and management with adequate analytical processes. Rather than implementing costly and complicated software applications within the organizations, BI can be used in accordance to the specific needs in a form of specialized cloud services.

The main advantages of the cloud BI systems (BIS) are related to the opportunity to introduce software as a service and infrastructure as a service in order to satisfy the increasing companies' demands of analytics. This elastic scalability gives strong benefits especially for small and medium size enterprises.

Over the last few years cloud solutions become the strongest trend of IT development. This trend naturally affects the BI solutions. BI vendors have already developed cloud BI services. Market expansion of cloud BI systems is growing rapidly.

2. BI systems in a Cloud

The development of business intelligent systems in a cloud environment can be considered as a radical change of the business procedures, mainly because of the fact that BI applications become accessible from anywhere at any time, which is not the traditional approach of BI systems.

Companies slowly but steadily migrate their BI applications from traditional models to cloud solutions. Cloud BI computing is a major topic of discussion since the beginning of 2010, and in 2011 many BI vendors develop strategies and solutions for introducing business intelligence in a cloud. In 2012 Cloud BI technologies increasingly penetrate the business environment and users are getting increasingly satisfied with the benefits. During 2013 this trend is confirmed in its progress and the Gartner projections expect the trend to continue and by 2015, 75% of the global scale 1000 companies to host the cloud BI solutions.

The types of Cloud services are specific for BI applications and the survey made by BI Leader Consulting (2013) demonstrates the trends for better acceptance of Cloud BI during the last year (Figure 1).

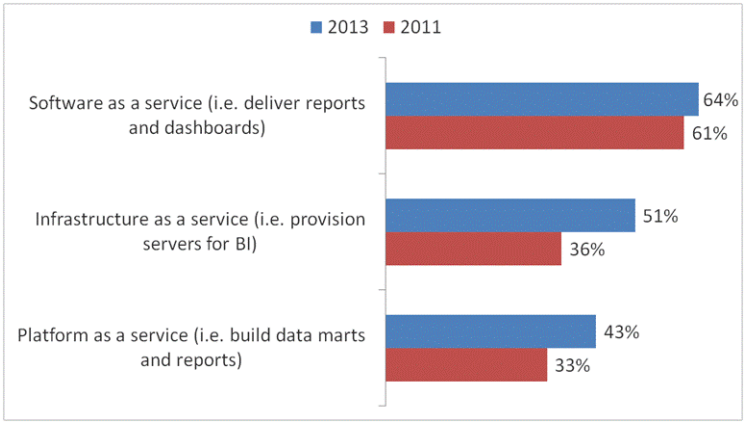


Figure 1: Cloud services for BI - 2011 and 2013 (Source BI Leader Consulting)

Monitoring the process of expanding the use of BI in a Cloud during the last year, it is very important to focus on the companies' size and their selection of the respected Cloud BI services (Figure 2). The figure data clearly shows that small companies do mainly introduce Software as a service (SaaS), while large and midsized companies are equally using SaaS and Infrastructure as a service (IaaS).

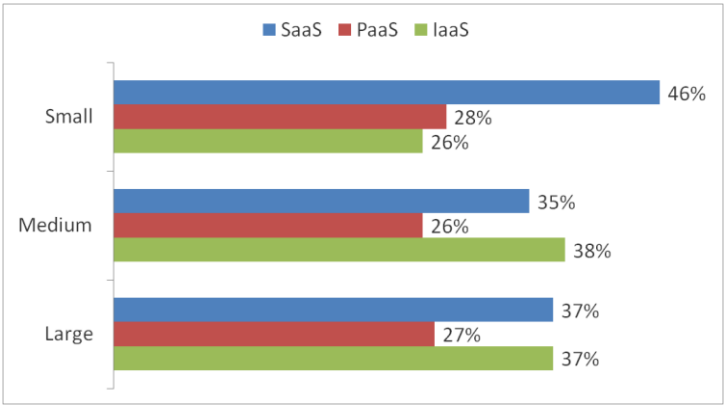


Figure 2: Types of Cloud BI services and companies size (Source BI Leader Consulting)

3. BI and Software as a Service

Software as a service (SaaS) is one of the most commonly used term in the field of cloud computing that allows all service users to receive the same application. Within a BI application this model allows all users to see the same file, but data to vary depending on the user's role and management rights in the organization.

SaaS Business Intelligent solutions allow companies to use the BI tools without in premises installation, additional maintenance and operational changes. Business users can mainly concentrate on BI analytics and reports.

There are certain concerns that should be considered before deciding the use of BI software as a service. Some of the services analytical tools may have limited functionality in comparison with the licensed BI software. Another concern that could prevent IT managers to choose a cloud BI solution relates to the confidentiality and risk to putt corporate data and information outside the organization.

However, one third of 1,364 IT managers surveyed by Gartner last year say they plan to use BI technology in a form of software as a service. More precisely - 27% of respondents say they already use or plan to use in the near future Cloud BI and to replace or upgrade the previously used BI software in their organization. Another 17% say they have already replaced or plan to replace in the near future licensed BI solutions with BI SaaS.

Main considerations in BI SaaS introduction:

• Data amount - whether the organization produces a sufficient amount of data to use SaaS BI solutions and obtain the requested information and analytics for effective management of decisions making. Different scalability of the SaaS solutions should be considered in order to get to the optimal level, depending on the amount of data of the particular organization.

• Technical evaluation – the choice for BI SaaS lacks experienced professionals to conduct the technical analysis and evaluation. Most of the BI users are managers at various levels in the organization, who have only business knowledge and for the BI SaaS technical analysis and evaluation they are supported by the Vendor IT professionals.

• Easy and fast solution for implementation - SaaS Business Intelligent solutions are available and accessible through a network environment and there is no need for installation of additional software, and purchase, installation, and maintenance of servers.

• Flexible payment - comparing to the traditional enterprise solutions, BI SaaS is flexible monthly/annual fees for the use of analytical tools. By using monthly payments,

users could regularly plan their analytical needs and project the needs for BI SaaS. Possibility for limited Free trials of SaaS enables users to use the technology on their own data before taking the decision for the software model.

• Getting a competitive advantage - the use of BI SaaS gives the organizations additional competitive advantages by helping them decide – who, when, where, what analytics to obtain in real time and meet the latest changes in both the organization and in its environment.

• SaaS BI complement other web based solutions - these are Salesforce, GoogleDocs, Google Analytics. SaaS BI solutions connect directly to online sources and all the necessary analyzes can be performed in a web browser. SaaS Business Intelligent tools are designed and created to complement the tools used by the organization.

For organizations that want to develop their own BI applications on their own data in order to provide their own customers, this model is not suitable. Such organizations could be better served by the other cloud services models.

4. BI and Infrastructure as a Service

Infrastructure as a service is a model that allows users to select the required infrastructure suitable for their business needs and to have a full control over the business intelligent applications they host and configure.

For organizations that need to develop their own BI applications, over their own data in order to serve their own customers, this model is an appropriate solution. In this way, the business intelligent application could be hosted and configured in the selected by the company environment with the preferred tools and instruments.

To realise this type of hosting, organizations should transfer to the cloud infrastructure their own applications that proves to appear a hard task for some users. According to the Gartner survey, there are five main approaches that organizations could select to transfer their applications in the cloud:

- The first option refers to the change of the host and reallocation of the applications to the different hardware environment and then reconfiguration of the infrastructure.
- The second method is called refractioning. Its purpose is to modify the source code of the application or its modules without affecting the application external behavior.
- The third option is a modification or extension of the existing code in order to provide a support to the cloud technologies.
- The fourth way is to build a new solution in a form of platform as a service. It refers to closing down the current applications and building new ones.
- The fifth option, similar like the fourth, refers to closing down the current solutions, but instead of developing new versions of them to search for available applications and implement them in the form of software as a service.

5. BI and Platform as a Service

Combining BI solution with the model platform as a service, allows users to implement BI applications from the hosted platform technologies. Some of the BI providers require the platform implementation to be based on a specific operating system like Windows.

Business intelligent PaaS applications typically support multiuser interfaces, which is implemented by providing a mechanism for managing competing applications, scalability, failures protection and ensuring high security levels. The architecture allows definition of

levels of trust in the relations between users, levels of security, access, program code sharing, logs navigation, user profiles, history of interactions and logs of applications usage.

There are several main types of platforms as a service that could support intelligent business systems:

- Additional development environment - such environments allow customization of existing software as a service application, and could be considered equivalent to the macro language that provides an environment for development and customization of existing packages of software applications. Often they require PaaS developers and end-users to purchase subscriptions for joint implementation and use of the two models.
- Individual environments for development – the individual PaaS environment for development of business intelligent applications does not include and is not dependant on technical, licensing or financial aspects of the other Cloud models SaaS and IaaS, as well as specific applications or Web services of software as service and are designed to provide a complete development environment.
- Environment for providing applications only - platform as a service from this type does not include an application development environment - development, debugging and testing as part of the service. Additional to the environment are supported offline tools for these purposes, such as Eclipse development environment. The limitations of the environment in the online mode are to maintain high levels of security in this model usage.
- An open platform as a service - this model platform as a service does not introduce a hosting plan, but provides open source software tools to develop and run applications. It allows users to deploy some applications written for Google App Engine on their own servers, providing access to databases via standard SQL or NoSql databases. Some of these open platforms allow developers to use any programming languages, database, operating system or server system that they prefer to upload the applications.

6. Challenges to BI in a Cloud

During the years of economic difficulties, cloud computing and business intelligence become increasingly important for companies to create and maintain competitive advantages. Combining BI and cloud computing can enable a variety of new business opportunities. However, there are still challenges in incorporating the two technologies.

In terms of operational, tactical, strategic BI and willingness of users to keep the business strongly in line with the economic deviations, the biggest challenges facing by Cloud BI are mainly:

- Data security for the sensitive data and distributing the analytics, unless a private cloud.
- Customization of the key performance indicators (KPI), which users would define and use - organizations have specific business processes and require monitoring of specific KPIs.
- Scalability of analytics to be handled by organizations with high-volume transactional data produced daily.
- Storage and handling of large volumes of data.
- Integration between cloud BI applications and data provided by the traditional systems.

- Many enterprises are not satisfied with physical, network and external security of the cloud computing platforms.
- New technology - cloud computing with BI is still a new segment and companies should anticipate and adapt to unforeseen and unexpected problems.
- Movement of data - organizations should be responsible to create and adopt new procedures for protecting their internal and confidential data that is getting exported outside of the organization.

7. Conclusion

During the recent years, two new technologies are steadily gaining strength and successfully penetrate various business environments. These two technologies are Business Intelligence and Cloud Computing.

Each of these technologies is strong enough to bring to organizations defined competitive advantages, increasing their effectiveness and support their business development.

Collaboration of the two technologies or so-called Cloud BI is turning to become a unique opportunity for the businesses to better analyze and understand the internal and external information and get competitive advantages. This collaboration is still in its early stages and requires a lot of research, analysis and elaborations in order to figure out lessons and create best practices.

References

1. Business intelligence in the cloud, http://sajim.co.za/index.php/
2. Cloud Hosting Opens Possibilities for Business Intelligence, http://www.thewhir.com/web-hosting-news/cloud-hosting
3. Dresner Advisory Study Shows Business Intelligence in the Cloud Increasingly Important Despite Security Aspects, http://sandhill.com/article/dresner-advisory-study-shows-business-intelligence
4. The report on the 2013 Cloud Business Intelligence Market Study, www.cloudbireport.com.
5. 2013: Year of Cloud BI – www.tdwi.org
6. Cloud-based Business Intelligence – A Market Study - www.biscorecard.com
7. Secure Cloud Business Intelligence - www.birtondemand.com
8. Business Intelligence Journal | Vol. 17, No. 3– www.tdwi.org
9. TDWI E-Book: Overcoming Obstacles with BI and Big Data–www.tdwi.org
10. The BI Revolution: Cloud BI Progress and Pitfalls - www.tdwi.org
11. Cloud business analytics – www.pentaho.com
12. Trends in Cloud BI adoption - http://www.insideanalysis.com
13. Cloud-based BI is a game-changer - http://www.infoworld.com
14. To be (to cloud) or not to be (not to cloud) BI - http://blogs.forrester.com/boris_evelson
15. Cloud BI – the data integration challenge - http://biblog.arcplan.com
16. BI Trends 2013 – Mobile BI, Collaborative BI, Cloud BI -http://biblog.arcplan.com

An Analysis of Life Cycle Costing Using Fuzzy Logic

Plamena Zlateva[1], Georgi Zabunov[2], Dimiter Velev[2]

[1] ISER - BAS, Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria
plamzlateva@abv.bg

[2] UNWE, 1700 Sofia, Bulgaria
georgi_zabunov@unwe.bg, dgvelev@unwe.bg

Abstract. Life Cycle Costing (LCC) can help organizations apply knowledge about uniting into one the past performance, future costs and investment risks. A hierarchical fuzzy logic system with three inputs and one output is designed. The system included two fuzzy logic subsystems. The proposed fuzzy logic system for risk estimation of the investing in improper facility can be successfully used in the implementation of Life Cycle Costing Analysis in making investment decisions.

Keywords. Life Cycle Costing, Fuzzy logic system, Investment risk estimation.

1. Introduction

The rapid development of information technology changes the way we think about real estate. More and more buildings are transformed from ordinary buildings in technologically complex facilities. The role of technical specialists in making complex investment decisions increases. There is a growing need for technical specialists and investment experts to speak the same language. The concept of Life Cycle Costing (LCC) provides such an opportunity.

Some would say that LCC is to help engineers think like economists and investors. This concept should be considered in more depth. It could be stated that the main purpose of LCC should be to help organizations apply knowledge about uniting into one the past performance, future costs and risks. This should be done in meaningful predictions about future costs of products, processes, and organization, and their associated business risks [1]. This is why the discussion of such important topics as sustainable cities, smart buildings, smart cities, urban security and many others, can not be done without the concept of LCC.

Considered in the most general terms, the Life Cycle Costing refers to the total cost of owning an asset over its entire life. Some researchers considered the concepts of Life Cycle Costing (LCC) and Whole Life Costing (WLC) as equivalent. We argue that there is difference between the two concepts, but it is not essential. The difference is due to different viewpoints.

The ISO 15686-5:2008(E) standard defines in detail the concept of Life Cycle Costing (LCC). Whole life costs include life cycle costs. For its part, life cycle costs include all costs such as design and building costs, operating costs, associated financing costs, depreciation, and disposal costs. However, whole-life cost also takes certain costs that are usually overlooked into account, such as environmental impact and social costs.

In other words, the concept of LCC is a look at cost through the eyes of the consumer or investor, and the concept of WLC through the eyes of society as a whole. The aforementioned ISO 15686-5:2008(E) standard contains as elements of WLC not only LCC

but also non-construction costs, income and externalities. Respectively, according to the same standard, the main elements of LCC are connected with construction, maintenance, operation, occupancy and end of life.

The purpose of this paper is to propose a fuzzy logic system for risk estimation of the investing in improper facility can be successfully used in the implementation of Life Cycle Costing Analysis in making investment decisions. The basic idea is to use the knowledge of experts. This fuzzy logic system can be incorporated into Web-based expert system.

2. Life Cycle Costing Concept

The LCC concept gives many advantages to organizations. Traditional cost management focuses only on some aspects of an organization's costs. It has lost its relevance [2]. A major obstacle to making appropriate investment decisions in the organization is the structure cost orientation. Structure orientation means that costs are categorized according to their types or structure. An example of this approach is a classification of costs like marketing costs, direct labor costs, depreciation, and so on. This way of thinking does not consider what kind of jobs have to be done. Following this approach, the management has to implement cost cutting in order to achieve cost effectiveness. The most important thing is that it will lead to cost cutting, starting with the largest costs. Such cost cutting can be ineffective, because one might cut costs for jobs that are very valuable to the organization rather than the intended idle capacity. As a result the company is going to reduce its ability to produce value and therefore will be much worse.

The only way to cope with this problem is by thinking in terms of processes. That means to implement so-called process orientation. The costs should be categorized according to the process elements like activities, tasks, jobs and so on. Working in this way, the manager is able to assess the full costs associated with a particular facility (labor, spare parts, depreciation, energy and so on). The point is that it is easier to measure the time the machine is idle, estimate the costs, and find out how to change this activity. Furthermore, the process orientation is a direct link to quality management. Cost of quality can be realistically estimated.

In short, the LCC concept entails three shifts: from a partial focus to holistic thinking; from structure orientation to process orientation; from cost allocation to cost tracing [1]. As a result, the LCC concept improves the management of risk and uncertainty in the organization.

The most valuable feature of LCC is that it allows the management of investment risk in complex construction projects. The foundation of effective risk management lies in the reasonable Life Cycle Cost Analyses (LCCA). LCCA is a process of evaluating the economic performance of a building over its entire life. Sometimes known as "whole cost accounting" or "total cost of ownership," LCCA balances initial monetary investment with the long-term expense of owning and operating the building [3]. The values for the costs resulting from the LCCA are the basis of measuring the expected return on the project, and hence the project risk. One very serious difficulty appears here. The life cycle of the building as a whole is a result of enduring life cycles of all its components. All these life cycles are of different duration. It is not possible to find a common cycle whose duration is a multiple of all the others. Therefore, the overall assessment of risk throughout the lifecycle of the building will consist of assessments of the risk of the individual components in their own life cycles. This can be illustrated by the following example [3, p. 14]

Annual Investment in Plant Assets – Subsystem Life Cycles

Subsystem Categories	Average Life Cycle
1a. Roofing – Tile	80 years
1b. Roofing – Metal, Concrete	50 years
1c. Roofing – Membrane, Built-up, Shingle, Bitumen, Foam	20 years
2a. Building Exteriors, Doors, and Windows (Hard)	80 years
2b. Building Exteriors (Soft)	20 years
3. Elevators and Conveying Systems	25 years
4. HVAC – Equipment and Controls	20 years
5. HVAC – Distribution Systems	40 years
6. Electrical Equipment	30 years
7. Plumbing Fixtures	30 years
8. Plumbing – Rough-in	50 years
9. Fire Protection Systems	40 years
10. Fire Detection Systems	20 years
11. Built-in Specialties and Equipment	25 years
12. Interior Finishes	15 years
Other Categories	
13. Foundations	Lifetime
14. Subgrade drainage and waterproofing	As needed
15. Vertical Elements	Lifetime
16. Horizontal Elements	Lifetime
17. Interior Partitions	As needed
18. Electrical – Rough-in	Lifetime
19. Site Preparation	Lifetime
Categories Included as Infrastructure	
20. Site Development – Softscape	Infrastructure
21. Site Development – Hardscape	Infrastructure
22. Site Development – Distribution	Infrastructure
23. Site Utilities	Infrastructure

It is obvious that if we apply this approach, risk assessment becomes enormously difficult. However, even that is not all. Let us not forget that LCC is a process, but not a single act. It incorporates many consecutive steps as: Defining the objective of the proposed LCC analysis; Preliminary identification of parameters and analysis requirements; Confirmation of project and facility requirements; Assembly of cost and performance data; Carry out analysis, iterating as required; Interpreting and reporting results. Seen from this

perspective, we need to look at ways to integrate into an overall assessment the scores for individual steps of the process.

3. Design of Fuzzy Logic System

Fuzzy logic systems are defined as intelligent systems for decision making, which are capable of dealing with uncertainty, inaccuracy and partial authenticity of the quantity and quality information available. Fuzzy logic systems ensure higher effectiveness and robustness of the managerial decisions that have been made. The effect of the fuzzy logic systems is analogous to the behavior of a group of highly qualified experts with experience in the area of research, which make informed decisions in conditions of uncertainty [4, 5].

Risk assessment for a project as a whole by evaluating the individual stages is based on the subjective knowledge of experts about the possible risks of a certain stage and the LCCA in the organization. Expert assessments are represented through linguistic variables (with values of little, average and big risk respectively), which by nature are quality not quantity variables. That naturally leads to the idea of assessing the risk using fuzzy logic.

Different approaches are possible. One of them is to develop a hierarchically fuzzy expert system, within which, for example, three quality criteria of analyzing the project risk (of a construction project) have been defined:

1. Risk of investing in a facility, which is based on inefficient technology - *K1*
2. Risk of investing in facilities for which new technologies emerge - *K2*
3. Risk of shortage of funding - *K3*

The above criteria are interpreted as linguistic variables, which are input for the fuzzy expert system.

The fuzzy logic system output is linguistic variable "Estimation of Risk of the investing in improper facility".

To construct the expert system it is assumed that the design and construction of a building and its facilities are crucial for the whole project. The first two risks (*K1* and *K2*) are related to the design stage, and the third (*K3*) to the stage of construction. Accordingly, at each of these stages corresponds to fuzzy subsystem, as hierarchical levels of the expert system (FS1 and FS2).

Therefore, a hierarchical expert system is used and it includes two levels and two fuzzy subsystems. Thus, the logical rules, which arise from the relationship of the three input variables, are distributed in two knowledge databases. That structure of the fuzzy system allows a more accurate representation and coding of expert knowledge, as well as an easier clarification of the rules. A generalized image of the two-level hierarchical system is shown in Figure 1, where FS 1 is the fuzzy subsystem 1 and FS 2 – the fuzzy subsystem 2.

It can be seen the first and second level of the hierarchical system under consideration includes one fuzzy subsystem. Each subsystem has two inputs and one output. This type of configuration reflects the impacts of the previous system elements to the next elements.

For the first fuzzy logic subsystem the inputs are the two linguistic variables (basic criteria - assessment of the possible risks to a given stage):

1. "Risk of investing in a facility, which is based on inefficient technology" – *K1*;
2. "Risk of investing in facilities for which new technologies emerge " – *K2*.

The output is the linguistic variable "Level of risk on the design stage" – *K4* (intermediate complex criterion).

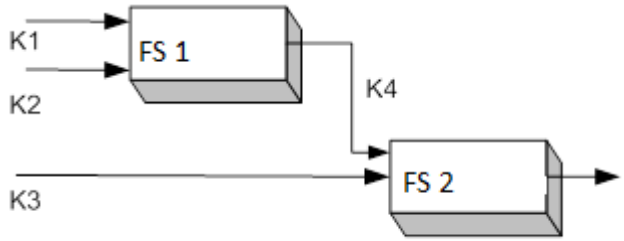


Figure 1. Two-level hierarchical fuzzy system

The fuzzy logic subsystem at the second hierarchical level has the following input variables (one intermediary complex criterion and one basic), as follows:

1. "Level of risk on the design stage"– *K4;*
2. "Risk of shortage of funding" – *K3*.

As the output of the fuzzy subsystem, the variable "*Estimation*" (Estimation of Risk of the investing in improper facility) is introduced. This also is the output for the whole fuzzy logic expert system. It is the final complex estimation of the project risk.

As it is known from the theory of the fuzzy sets, linguistic variables can take different quantity or quality meanings (little-big; low-middle-high, etc.). In fuzzy logic systems these variables are considered to be sets (therms) with a certain degree of membership (membership function) to the specific value. The membership functions are defined with different forms as triangular, trapezoidal, etc.

In this case the input variables (the three basic criteria) and the intermediate complex criterion are represented by means of three fuzzy sets for the two subsystems: "Little" ("Little"- L), "Middle" ("Middle"- M), "Big" ("Big"- B) level.

The output variables of the hierarchical fuzzy logic system is defined with five sets: "Very Little" ("Very Little"- VL), "Little" ("Little"- L), "Middle" ("Middle"- M), "Big" ("Big"- B) and "Very Big" ("Very Big"- VB).

In this paper, all linguistic variables are assigned with normal triangular functions of appurtenance. The variables - *K1, K2, K3, K4* are changed within the interval [0, 10], and the output variable ("*Estimation*") – in the interval [0, 100]. The form of the linguistic variables is shown in Figures 2 and 3.

The rules for making a conclusion in knowledge data bases are defined by means of *"If - Then"* logical forms. Nine rules are included in each of the fuzzy subsystems. Some of these rules are defined as follows:

IF [*K*1 is *Little*] and [*K2* is *Little*] THEN [*K4* is *Little*];
IF [*K1* is *Little*] and [*K2* is *Big*] THEN [*K4* is *Middle*];
IF [*K1* is *Middle*] and [*K2* is *Little*] THEN [*K4* is *Little*];

IF [*K4* is *Little*] and [*K3* is *Big*] THEN [*Estimation* is *Middle*];
IF [*K4* is *Little*] and [*K3* is *Middle*] THEN [*Estimation* is Little];
IF [*K4* is *Middle*] and [K3 is *Big*] THEN [*Estimation* is *Big*];
IF [*K4* is *Big*] and [*K3* is *Little*] THEN [*Estimation* is *Middle*].

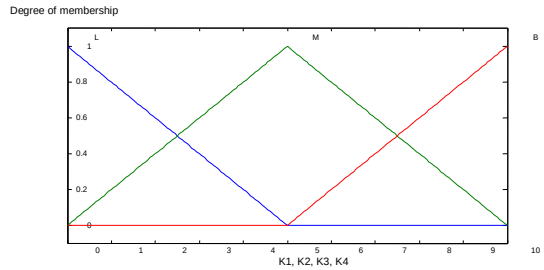


Figure 2. Form of the input linguistic variables (K1, K2, K3, K4)

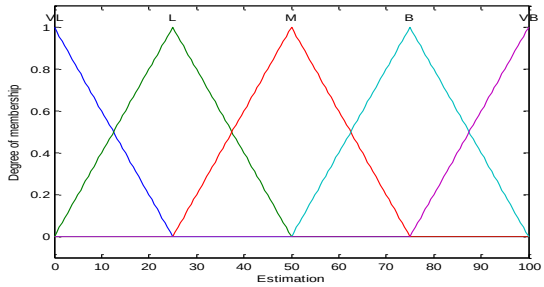


Figure 3. Form of the output linguistic variable "Estimation"

The hierarchical fuzzy logic system for complex risk estimation of construction project is projected in *Matlab* environment, using Simulink and Fuzzy Logic Toolbox. The two subsystems are designed in "Mamdani" variant of fuzzy systems. Classical max/min procedures of processing the rules and method of defuzzying – center of gravity are used.

The output surfaces for the first and the second fuzzy subsystems in three dimensional space, with coordinates (*K1, K2, K4*) and (*K4, K3, Estimation*) respectively, are shown in Figure 4.

The initial fixed values of the input variables, which are subject to fuzziness, can be obtained on the basis of analysis and appropriate statistical processing of data collected through questionnaires completed by experts.

The particular value at the output of the hierarchical fuzzy logic system is the final complex estimate of the investment risk, associated with a certain construction project. Therefore, the output value of the fuzzy logic system can play the role of an indicator and to facilitate the making of more informed managerial decisions.

4. Conclusion

A hierarchical fuzzy logic system with three inputs and one output is designed. The system included two fuzzy logic subsystems. The proposed fuzzy logic system for risk estimation of the investing in improper facility can be successfully used in the implementation of Life Cycle Costing Analysis in making investment decisions. The basic idea is to use the knowledge of experts. On a later stage, this fuzzy logic system can be incorporated into Web-based integrated expert system.

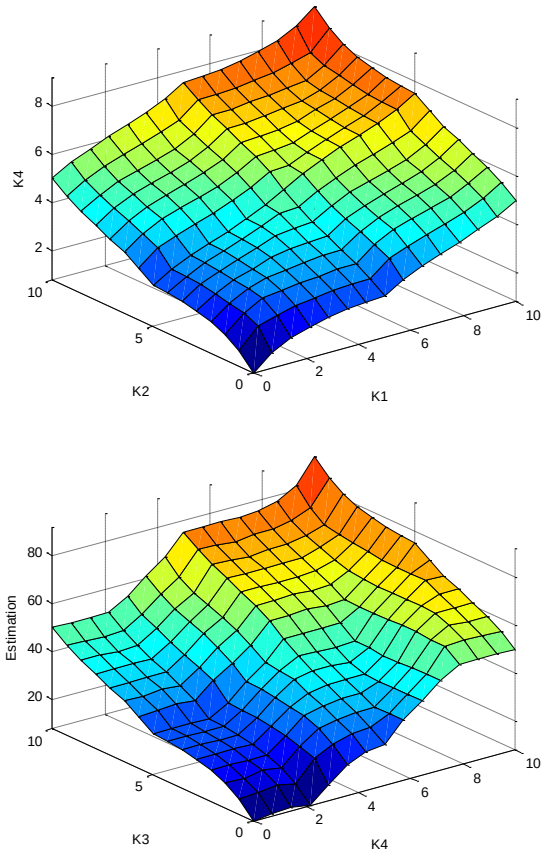


Figure 4. The output surfaces for the first and the second fuzzy logic subsystems

References

1.Emblemsveg J., *Life Cycle Costing*, Hoboken, New Jersey: John Wiley & Sons, Inc., 2003.
2.Johnson H.T., R.S. Kaplan, *Relevance Lost: The Rise and Fall of Management Accounting*. Boston, MA: Harvard Business School Press, 1987.
3.Guidelines for Life Cycle Cost Analyses, Stanford University, Land and Buildings, October 2005. (http://lbre.stanford.edu/sites/all/lbre-shared/files/docs_public/LCCA121405.pdf)
4.Zadeh, L. A. et al., *Fuzzy Sets, Fuzzy Logic, Fuzzy Systems*, World Scientific Press, 1996
5.Zlateva P., D. Velev, G. Zabunov, A Model for Fuzzy Logic Assessment of Real Estate Investment Risks, *Advances in Intelligent and Soft Computing*, vol. 101, 2011, pp 89-93.

Multivariate Analysis of Steel Alloys Components and Characteristics Using Copula Approach

S. Popova[1], P. Koprinkova-Hristova[2], P. Zlateva[1], A.Toncheva[3]

[1] ISER - BAS, Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria
Popova_Silvia2000@yahoo.com, plamzlateva@abv.bg

[2] IICT - BAS, Acad. G. Bonchev Str., Bl. 25A, 1113 Sofia, Bulgaria
pkoprinkova@bas.bg

[3]SU Kliment Ohridski, Car Osvoboditel 15
tontchev@gmail.com

Abstract. Production of high strenght steel alloys is of big importance for the modern metallurgy.The relationship between steel alloys components and characteristics is investigated. Describing these parameters as multivariate random variables, temporal nonlinear dependencies between them are investigated applying Kendall's tau and Spearman's rho rank correlation coefficients and copula approach. Conducting a multivariate analysis of the available data, possibilities and limitations of copulas to establish the nature of the interdependence between steel alloys components and characteristics are discussed. Because Kendall's tau and Spearman's rho rank correlation coefficients are monotonous, we make conclusions which components and characteristics increase or decrease together. The calculations are performed using Matlab software.

Keywords. copula, steel alloys composition, steel alloys characteristic

1. Introduction

Production of high strenght steel alloys is of big importance for the modern metallurgy. The main aim is to obtain high quality materials reducing quantitiy of used expensive compounds. Another ecollogically motivated reason to work on steels composition optimization is obtaining of ultra high strenght steels (from the lower right part of the Figure 1) that will allow production of lighter machines – goal that will result finally in reduction of CO_2 emmisions in the atmosphere.

In any case the identified high dimentional MIMO models are used further for solving of multicriteria optimization task. The multiple objective mathematical programming (MOMP) approach was usually applied in case of regression models. In any case of neural network models another useful characteristic of them can be applied in optimization procedures: neural networks are able to accumulate knowledge by experience using training procedures that are in fact optimization of neural network weights with respect to error at neural network output [2,9];thus neural network training procedures offer a common approach to optimization tasks in process optimization and control applications [2,3,8,13,14].

In the present paper the relationship between steel alloys components and characteristics is investigated. Describing these parameters as multivariate random variables, temporal nonlinear dependencies between them are investigated applying Kendall's tau and Spearman's rho rank correlation coefficients and copula approach. Conducting a multivariate analysis of the available data, possibilities and limitations of copulas to establish the nature

of the interdependence between steel alloys components and characteristics are discussed. The investigations can be used in practice in design of steel allows when the predictions are in the intervals

2. Study of the Interdependence of Multivariate Random Variables by Copula Approach

In the late sixties of last century Sklar [12] introduced a copula concept to separate the effect of dependence from the effect of marginal distributions in a joint distribution. The copula functions provide a natural way to study and measure linear and nonlinear dependencies between multiple random variables. Recently, the copula approach is widely used in various fields. The procedures of copula approach are presented in details in statistical literature [1, 4, 5,7, 10, 11]. Here, only short description of the analysis of multivariate random variables is presented.

The linear correlation or Pearson's correlation is most frequently used in practice as a measure of dependence between two random variables X and Y. However, since the linear correlation is not a copula-based measure of dependence, it can often be quite misleading and should not be taken as the canonical dependence measure. Two important nonparametric measure of statistical dependence (concordance) are Kendall's tau and Spearman's rho correlation coefficients. They measure the degree of monotonic dependence, while Pearson's correlation coefficient indicates only linear dependence.

According to Scarsini [10] the dependence is a matter of association of the X and Y variables along any (measurable) function, i.e., the more X and Y tend to cluster around the graph of a function, either $Y = F(X)$ or $X = G(Y)$, the more they are dependent. The minimum dependence, as well as the minimum of monotone dependence, corresponds to independence between two variables. Concordance takes into account the kind of monotonicity (whether increasing or decreasing), so that the maximum of concordance is attained when a strictly monotone increasing relation exists between the variables, and the minimum of concordance (perfect discordance) is attained when a relationship exists that is strictly monotone decreasing.

Lancaster [5] adds that "a measure of dependence indicates in some defined way, how closely X and Y are related, with extremes at mutual independence and (monotone) dependence." Two key differences between "dependence" and "concordance" can be outlined [5, 10]:

- Opposed to the dependence measure, which reach their minimum, when the random variables X and Y are independent, the concordance measure reaches it minimum when X and Y are countermonotone;
- Independence is a sufficient, but not necessary condition for the concordance measure to be equals to zero.

Without a strict mathematical definition, the concordance between two random variables can be described as capturing the fact that the probability of occurrence of large (small) values of both X and Y is high. Alternatively, the probability of occurrence of high values of X, together with small values of Y is low. Nelsen [7] demonstrates that if X and Y are continuous random variables whose copula is C(X, Y), then the two popular copula-based concordance measures: the rank correlations Kendall's tau and Spearman's rho satisfy the properties for a measure of concordance. For more details about Kendall's tau and Spearman's rho and their estimators we refer to Kendall and Stuart [4], Nelsen [7] and Schmidt [11].

To investigate interdependences between steel alloys components and characteristics the Kendal's tau and Spearman's rho correlation coefficients are calculated. Then, cumulative

distribution functions and probability density functions are estimated using Frank copula. This Archimedean copula is most appropriate for data that exhibit weak tail dependence [1].

3. Experimental investigations. Results and analyses

Data set

In our investigation we've used a data base containing information about 99 steel alloys available at http://www.splav.kharkov.com/choose_type.php. There are given concentrations of eleven alloyng elements: C, Si, Mn, Ni, B, S, P, Cr, Cu, Mo, V included in the steels and values of six steel strenght characteristics tested after thermal treatment of the steels as follows: *Rm* – tensile strenght; *Re* – yield strenght; *A* –elongation; *Z* – reduction of area; *KCU* – impact strenght; *HB* – Brinell hardness

The multivariate correlation structure between variables and stochastic interdependences between them is assessed using copula approach. The calculations are performed on PC using Matlab software.

Dependences between steel alloys components and characteristics

For all pairs of random variables correlation coefficients, Pearson, Kendall's tau and Spearman's rho are calculated and the results are given in Table 1.

Graphical representation of copula approach application

There are highest positive (correlations) relationships between Ni and HB. (Figure 2) It is clearly that a strong negative correlation between S/P and HB exists. (Figure 3). Figure 2 and Figure 3 can be further analyzed jointly with the respective copula cumulative density functions presented in first columns and row 1. A straightforward way to gain understanding of the $[0,1]^3$ visualization of the copula probability density function is by analyzing the intensity and location in the [0,1] plane of the level sets of its contour plots, which are also presented on Figure 2 and 2 in second column.(row 2). The Frank copula exhibits asymptotically lower tail dependence, which is proven by the increasing intensity of the level lines of the contour plot of the copula probability density function, as they approach the [0,0] zone of the contour plot graph on Figures 2 and 3, Figures 4 and 5 present copulas for highest and lowest rank correlation from table 1, respectively.

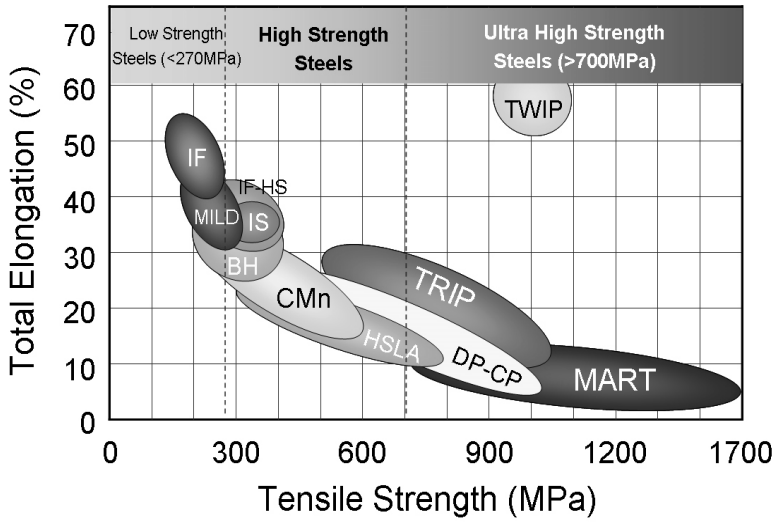


Figure 1. Steel's types dependence on their mechanical characteristics

Table 1 Pairs' correlation coefficients (CC)

Pairs's correlation coefficients	Pearson CC	Kendal rank CC	Frank copula dependence parameterKendal	Spearman rank CC	Frank copula dependence parameter Spearman
C-Rm	-0.022	0.016	0.144	0.002	0.012
Si- Rm	-0.008	-0.044	-0.401	-0.055	-0.328
Mn-Rm	-0.21	-0.138	-1.262	-0.191	-1.165
Ni-Rm	0.300	0.217	2.032	0.299	1.874
B-Rm	0.108	0.078	0.706	0.091	0.55
S/P-Rm	-0.325	-0.243	-2.3	-0.294	-1.839
Cr-Rm	0.294	0.225	2.114	0.318	2.009
Cu-Rm	-0.006	-0.035	-0.317	-0.038	-0.23
Mo-Rm	0.059	0.177	1.638	0.218	1.338
V-Rm	-0.081	-0.033	-0.293	-0.041	-0.248
C-Re	0.012	0.057	0.514	0.063	0.381
Si- Re	-0.009	-0.047	-0.42	-0.055	-0.331
Mn-Re	-0.252	-0.156	-1.433	-0.213	-1.306
Ni-Re	0.302	0.191	1.770	0.265	1.648
B-Re	0.109	0.097	0.881	0.117	0.708
S/P-Re	-0.355	-0.272	-2.612	-0.332	-2.104
Cr-Re	0.375	0.278	2.677	0.388	2.515
Cu-Re	-0.08	-0.095	-0.858	-0.113	-0.682
Mo-Re	0.161	0.261	2.487	0.326	2.067
V-Re	-0.04	0.037	0.336	0.046	0.276
C-A	-0.054	0.016	0.146	0.027	0.159

Si- A	0.238	0.096	0.87	0.115	0.697
Mn-A	0.062	0.005	0.046	0.006	0.034
Ni-A	0.219	-0.165	-1.152	-0.232	-1.437
B-A	0.149	-0.034	-0.307	-0.038	-0.229
S/P-A	0.162	0.1	0.91	0.116	0.699
Cr-A	-0.224	-0.098	-0.892	-0.135	-0.82
Cu-A	0.236	0.16	1.466	0.185	1.131
Mo-A	-0.117	-0.121	-1.099	-0.144	-0.875
V-A	0.019	0.153	1.402	0.183	1.116
C-Z	-0.385	-0.285	-2.749	-0.378	-2.445
Si- Z	0.231	0.106	0.961	0.122	0.736
Mn-Z	-0.116	-0.159	-1.454	-0.209	-1.28
Ni-Z	-0.076	-0.063	-0.568	-0.083	-0.499
B-Z	0.224	0.129	1.174	0.145	0.878
S/P-Z	0.083	0.053	0.476	0.059	0.357
Cr-Z	-0.047	-0.045	-0.407	-0.06	-0.358
Cu-Z	0.126	0.115	1.046	0.131	0.791
Mo-Z	-0.27	-0.132	-1.206	-0.155	-0941
V-Z	-0.175	-0.021	-0.189	-0.025	-0.148
Pairs's correlation coefficients	**Pearson CC**	**Kendal rank CC**	**Frank copula dependence parameter Kendal**	**Spearman rank CC**	**Frank copula dependence parameter Spearman**
C-Kcu	-0.376	-0.267	-2.549	-0.372	-2.401
Si- Kcu	0.099	-0.031	-0.279	-0.035	-0.211
Mn- Kcu	-0.047	-0.192	-1.781	-0.252	-1.563
Ni- Kcu	0.1	0.086	0.781	0.122	0.739
B- Kcu	0.037	0.03	0.274	0.035	0.209
S/P- Kcu	-0.098	-0.099	-0.896	-0.116	-0.7
Cr- Kcu	0.017	0.005	0.042	-0.004	-0.021
Cu- Kcu	0.057	0.013	0.12	0.015	0.091
Mo- Kcu	-0.184	-0.002	-0.114	-0.009	-0.055
V- Kcu	0.09	0.024	0.219	0.029	0.172
C-HB	0.106	0.129	1.172	1.188	1.147
Si- HB	0.001	0.045	0.404	0.048	0.286
Mn- HB	-0.243	-0.319	-3.138	-0.431	-2.851
Ni- HB	0.355	0.424	4.479	0.542	3.843
B- HB	0.249	-0.135	-1.229	-0.16	-0.972
S/P- HB	-0.534	-0.589	-7.664	-0.681	-5.529
Cr- HB	0.305	0.382	3.295	0.456	3.065
Cu- HB	-0.007	-0.099	-0.9	-0.117	-0.705
Mo- HB	0.239	0.414	4.359	0.495	3.396
V- HB	0.148	0.236	2.23	0.282	1.757

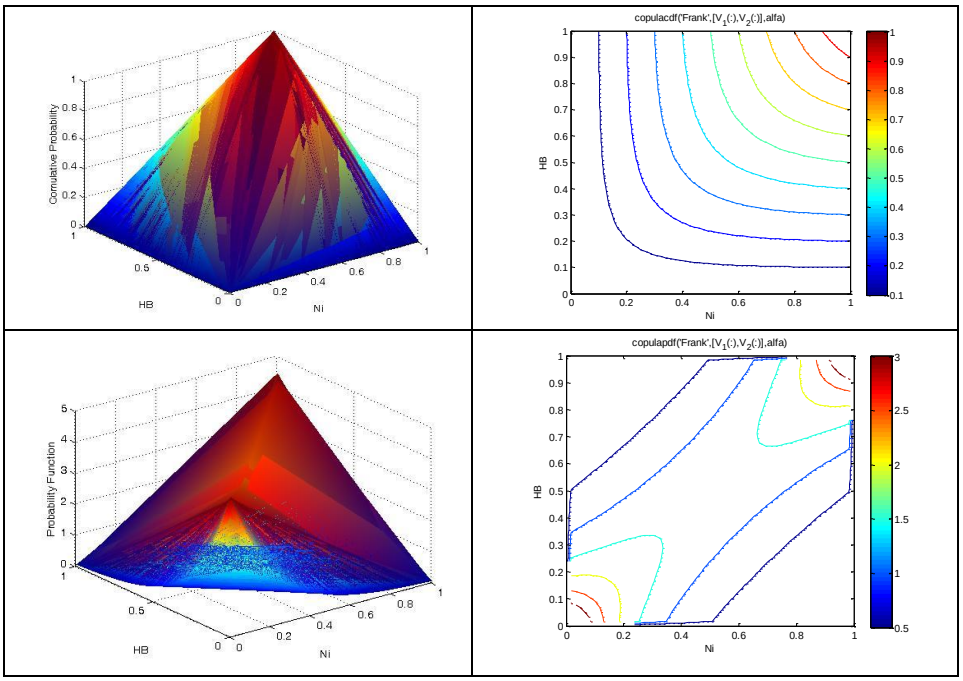


Figure 2.The cumulative function (column 1) and projection of the cumulative functions (column 2), probability density function and its projection (2 row) of Frank copula

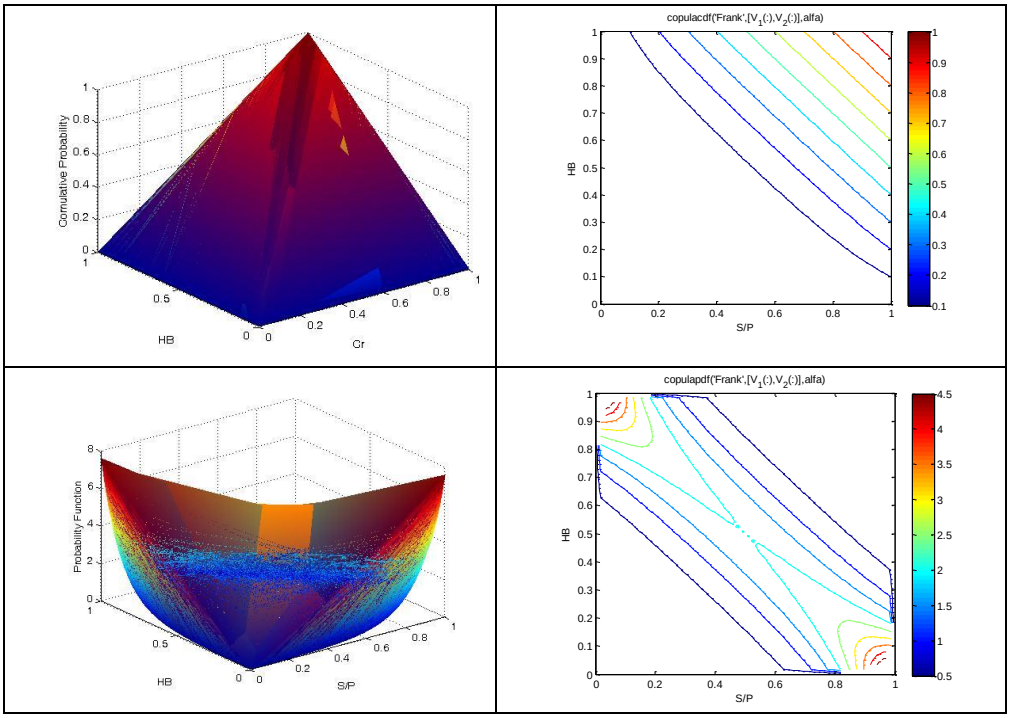


Figure 3. The perspective plot of cumulative density functions (column 1) and projections of the cumulative density function (column 2) ,probability density function and its projections(row2)of Frank copula, applied to the multivariate random values of steel alloys component S/P and characteristics HB

4. Conclusion

From the table and figures can be concluding that interrelationship between Ni and HB, followed by the relationship with precipitations is strongest. Because of the rank correlations Kendall's tau and Spearman's rho are measures of concordance the above results shows that [Ni, Rm], [Cr,Rm], [Ni, Re], [Cr,Re] [Si, A], [Cu, A], [B, Z], [Ni, HB], [Cr, HB], [Mo HB] values increase or decrease together.

The negative concordance is established between [S/P Rm], [S/P Re], [Ni A], [Cr A], [C Z], [Mo Z], [C Kcu], [S/P HB].These relations show that when the contents of component increase the characteristic decrease.

On the basis of the performed studies and analyses of the obtained results we can conclude that the copula approach allows reliable estimation of multiple correlations between the steel alloys components and characteristics. The investigations can be used in practice in design of steel allows when the predictions are in the intervals.

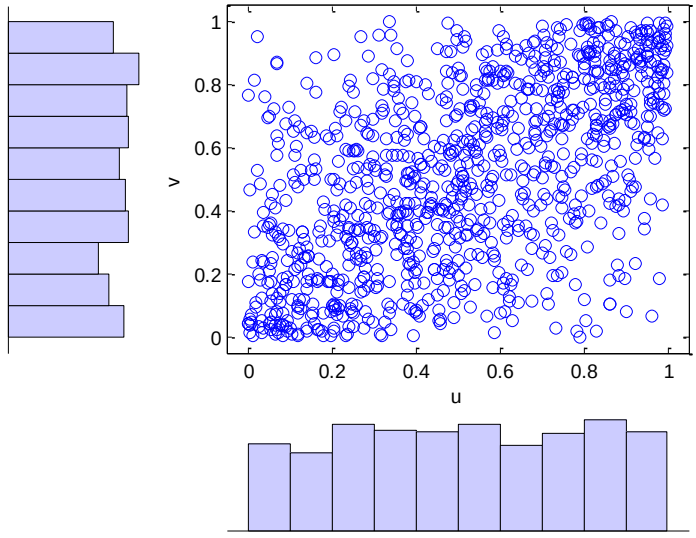


Figure 4- Frank copula, applied to the multivariate random values steel alloys component Ni and characteristics HB

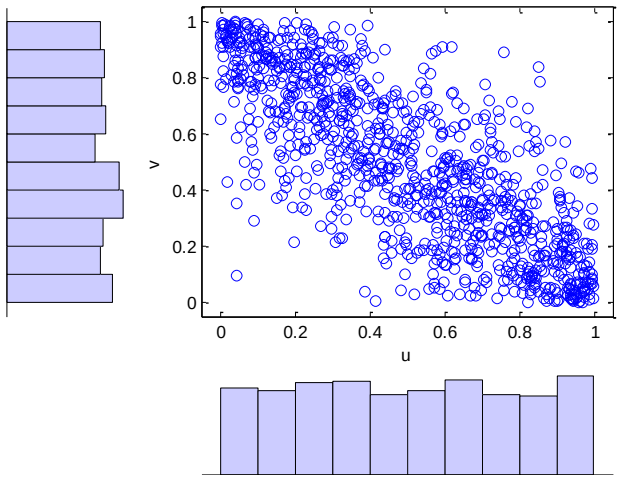


Figure 5- Frank copula, applied to the multivariate random values of steel alloys component S/P and characteristics HB

Acknowledgements. This work was partially supported by the Bulgarian National Science Fund under the Project No DDVU 02/11 "Characteristic modeling and composition optimization of iron-base alloys used in machine-building industry"

References

1. Cherubini, U., E. Luciano, and W. Vecchiato (2004) Copula methods in finance, John Wiley & Sons, 293 p.
2. D.E. Rumelhart, J.L. McClelland, Parallel Distributed Processing, vol. 1, MIT Press, Cambridge, MA, 1986.
3. D.H. Nguyen, B. Widrow, Neural networks for self-learning control systems, Int. J. Control, 54, No6 (1991), 1439-1451.
4. Kendall Maurice G and Alan Stuart (1979) The Advanced Theory of Statistics, Vol. 2 Inference and Relationship, 4th ed., Oxford University Press New York.
5. Lancaster, HO. (1982) Dependence, Measures and indices of dependence. In: Encyclopedia of Statistical Sciences, Volume 2, S. Kotz and NL Johnson (eds.), pp. 334–339. John Wiley and Sons, New York.
6. M.F. Anjum, I. Tasadduq, K. Al-Sultan, Response surface methodology: A neural network approach, European Journal of Operational Research, 101(1997), 65-73.
7. Nelsen, R. B. (2006). An introduction to copulas, Springer Series in Statistics, 2nd edition. New York: Springer, 269 p.
8. P.J. Werbos, An overview of neural networks for control, IEEE Control Systems (1991), 40-41.
9. P.J. Werbos, Backpropagation through time: What it does and how to do it, Proceedings of the IEEE, vol.78, No10 (1990), 1550-1560.
10. Scarsini, M. (1984). On measures of concordance. Stochastica. VIII, 201–218.
11. Schmidt, T. (2007). Coping with copulas. In Copulas - From Theory to Applications in Finance (J. Rank, ed.) 3–34. Risk Books, London.
12. Sklar, A. (1959) Fonctions de r´epartition `a n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Universit´e de Paris 8, 229–231.
13. Tontchev N., Z. Cekerevac Approach and Application in Multicriteria Decision Support in the Field of Materials Science. MEST Journal, (MEST) 2, no. 1 (01 2014): 18-29.
14. Tontchev N., M. Ivanov, Modeling and Optimization of the Composition of Iron – Based Alloys By Approximation with Neural Models and Genetic Optimization Algorithm, FBIM Transactions, 15 01, 2(1), pp. 1-12.

Cryptography Based on Finite Automata – a Brief Overview

Blagoj Ristevski[1], Violeta Manevska[2]
Faculty of Administration and Information Systems Management – Bitola
University St Kliment Ohridski Bitola, Republic of Macedonia

[1] blagoj.ristevski@uklo.edu.mk, [2] violeta.manevska@uklo.edu.mk

Abstract. In the past few decades, with advances in information and communication technology, the data sharing and exchanging by Internet is expanding rapidly. This growth is accompanied by data and information vulnerability, which are shared and exchanged. In order to become communication more secure, confidential and reliable there is a strong need to protect data and information broadened by computer networks and Internet. Therefore, in these few decades numerous modern encryption algorithms and systems are introduced such as DES, AES, IDEA, triple DES, cryptographic hash functions, RSA, ElGamal, elliptic curve, hyper-elliptic curve etc. In this paper, possibilities of application of automata theory to cryptography are briefly surveyed. The automata theory is very good basis for design of modern cryptosystems. Some of them are based on finite automata, whereas others are based on cellular automata. Additionally, the finite automata are described and their specific types. The concepts of invertibility, weakly invertible finite automata with delay τ are presented. To find an input sequence from an output sequence, searching algorithms are needed. In cryptography, besides finite automata, cellular automata can be applied. Public key cryptosystems based on finite automata can be applied on encryption and implementation of digital signatures. In addition, other applications of cryptography are exemplified. The end of the paper provides concluding remarks and further work toward development of new encryption techniques.

Keywords: cryptography, finite automata, public-key cryptography, cryptanalysis, digital signature.

1. Introduction

Cryptography is an interdisciplinary science whose bases are computer science, electrical engineering and mathematics. Automata theory is a mathematical theory that provides very good basis for developing cryptosystems. Some of them use finite automata, whereas others are based on cellular automata.

The rest of this paper is organized as follows. Section 2 describes the finite automata and their specific types of finite automata: autonomous, invertible, invertible with delay τ and compound. The concepts of invertibility, weakly invertible finite automata with delay τ are presented, too. To find an input sequence from an output sequence, searching algorithms are needed. Besides finite automata, the cellular automata can be applied in cryptography. In the following section, symmetric-key and public-key cryptosystems, as well as digital signatures are described. In addition, other applications of cryptography are exemplified. The final section of the paper provides concluding remarks and further work toward development of new encryption techniques.

Finite automata

Among the first applications of computer sciences was in cryptography. After inventing of Turing machine by Alan Turing in 1936, at that time the first computers was built to decrypt real cryptosystems for military purposes.

Finite automata as an abstraction of discrete and digital systems with finite "memory", are consisted of an input, an output and an internal state. The input, output and internal state are discrete. The input and the internal state at the current moment uniquely determine the current state of the output and the next time moment internal state. Finite automata transform an input sequence to an output sequence with the same length. They can be used as encoders to cipher the messages, based on the property of invertibility.

Let *X*, *Y* and *S* are nonempty finite sets, δ is a single-valued mapping from $S\times X$ to *S*, and λ is a single-valued mapping from $S\times X$ to *Y*. *X*, *Y* and *S* are called input alphabet, the output alphabet and the state alphabet of the finite automaton, respectively, whereas δ and λ are called next-state function and the output function of the finite automaton, respectively. Then, the quintuple $\langle X,Y,S,\delta,\lambda\rangle$ is called a finite automaton 3.

Let $M=\langle X,Y,S,\delta,\lambda\rangle$ is a finite automaton. If for any $s\in S$, $\delta(s,x)$ and $\lambda(s,x)$ do not depend on *x*, then *M* is **autonomous** finite automaton.

A finite automaton $M=\langle X,Y,S,\delta,\lambda\rangle$ is **invertible**, if for any $s,s'\in S$ and any $\alpha,\alpha'\in X^{\omega}$, $\lambda(s,\alpha)=\lambda(s',\alpha'),\alpha'\in X^{\omega}$ lieds to α=α'.

A finite automaton $M=\langle X,Y,S,\delta,\lambda\rangle$ is **invertible with delay τ** $(\tau\in N)$ if for any $s\in S$ and any $x_i\in X, i=0,1,\ldots,\tau,$ where $\tau\in N_0$, x_0 can be uniquely determined by $\lambda(s,x_0,\ldots,x_\tau)$. That means that for any $s,s'\in S$ and any $x_i,x_i'\in X, i=0,1,\ldots,\tau,$ $\lambda(s,x_0,\ldots,x_\tau)=\lambda(s',x_0',\ldots,x_\tau')$ leads to $x_0=x_0'$ 3.

Let $M=\langle X,Y,S,\delta,\lambda\rangle$ is a finite automaton. Finding of an input sequence from an output sequence is using by exhaustive search algorithms. An autonomous finite automaton can be regarded as a sequence generator 31.

Mathematically, a **cryptosystem** can be considered as a transformations family $\{f_k, k\in K\}$ that depends on a parameter (key) *k*, where f_k is a cryptographic transformation and *K* is a space of the keys. This transformation is an injective mapping from the plaintext space *P* to the space of ciphertexts *C*. The sender sends a plaintext message α to the receiver through an insecure channel. The sender first encrypts the message α applying f_k to the message, and then, the obtained result $f_k(\alpha)$, which is the cyphertext sent through the channel. The receiver decrypts the cyphertext by using an inverse transformation f_k^{-1} to obtain the original plaintext. When the sender and receiver share the same key *k*, the cryptosystem is called one key cryptosystem.

Considering the mathematical nature of finite automata and their implementation, they can be regarded as cryptosystems whose plaintext and cyphertext space are consisted of all words over particular finite sets, where the transformation $f_k(\alpha)$ corresponds to $\lambda(k,\alpha)$, λ is the output function of particular weakly invertible finite automaton, whereas the key space *K* is a seat of weakly invertible finite automata and their initial states.

Since development of the public-key cryptosystems, many cryptosystems are proposed in order to strengthen data and information security. From finite automata theory

viewpoint, the public key can be considered as a compound finite automaton of n+1 finite automata and states, where n+1>1. The private key is the n+1 weak inverse finite automata of them and states. Until now, there is no feasible inversion algorithm for the compound finite automaton, except the case when its decomposition is known 1.

The methods to construct an invertible finite automaton and find one of its inverses have two main applications: construction of the pairs of keys necessary for encryption, decryption and signature, and additionally, they can be applied to attack the actual finite automata-based cryptographic systems 8.

Besides application of finite automata, cellular automata can be applied in cryptography, which is described in 5 6 7 9 10. In 2 4, cellular automata are applied to generate a pseudo-random numbers sequence, which is used in the encryption.

2. Cryptography

Communication with classified secret ciphers was commonly used in diplomacy, during wars, for individual or company's privacy. It means that some people/companies want to be protected against their competition; especially they want to protect their sensitive information, benefits and resources against competitors, suspicious, curious or intelligence environment. That has led to emerge a need of encryption of the data, which is main subject of **cryptography**. Almost simultaneously with the cryptography, **cryptanalysis** has been developed. Cryptography is accompanied by cryptanalysis, whose aim is inverse of the aim of cryptography – to break the ciphers and codes. Cryptanalysis is used to break security of cryptosystems, hence to access to the encrypted messages, even if the key is unknown. Both, cryptography and cryptanalysis build the particular science called **cryptology**.

The **keys** are a shared secret between two or more parties that can be used to maintain a private communication channel. Depending on the encryption methods in which the sender and the receiver share the same key, or they use public key for encryption and secret key for encryption, cryptography is divided into **symmetric-key** and **public-key cryptography**, respectively.

In symmetric-key cryptography, the keys may be identical or they can be obtained by simple transformation from the other keys. The most popular and most commonly used symmetric-key cryptosystems are: *DES* (Data Encryption Algorithm), *AES* (Advanced Encryption Standard), *IDEA* (International Data Encryption Algorithm), *3DES* (Triple Data Encryption Algorithm), *Twofish*, *Blowfish*, *CAST5*, and *RC4*.

Asymmetric encryption is a form of cryptography, where one key is used to cipher the message and another to decipher. There is a mathematical relation between these two keys. If it is possible to make one of these keys public and still maintain the algorithm security, then the system is called public-key cryptosystem.

Public cryptosystem is a pair of algorithms $\{(E_k, D_k), k \in K\}$ for which the following is valid: for any $k \in K$, D_k and E_k should be inverse each other; D_k and E_k are easily to compute. It is infeasible to derive an easily computed algorithm to D_k from E_k and it is feasible to compute the pair of E_k and D_k.

The most commonly used public-key cryptosystems are: P*ohlig-Hellman*, *RSA*, *ElGamal*, *elliptic curve cryptography* (ECC) and *Diffie-Hellman* algorithm. Elliptic curve cryptography offers the benefit of smaller key sizes, thus smaller memory and CPU are required. The Diffie-Hellman, ElGamal, digital signatures, and the Digital Signature Algorithm (DSA) can all be implemented in ECC 11. This makes ECC a very suitable algorithm for wireless devices with limited bandwidth and processing power.

It is very important to insure that persons who are not authorized do not modify messages. The mechanism, which insures non-modification of the transmitted data from sender to receiver, even when it is stored, is called integrity. The most commonly used integrity mechanisms are: *Message Digest 5* (MD5), *Secure Hash Standards* (SHA), *Message Authentication Codes* (MACs) and *Keyed-Hash Message Authentication Codes* (HMAC) 11.

Public-key cryptosystems are very slowly and they are primarily used to encipher short messages. If a message is enciphered with the sender's private key and the receiver is able to decipher with the public key of the sender, then the sender cannot deny that the sender sent the message, because, the sender is the exclusive one with that private key. In order to provide authentication, non-repudiation, and integrity, **digital signatures** are used 11. Digital signature is made by taking the message's hash and encrypting it with the private key of the sender. By encryption, the inbuilt integrity is provided because if a ciphertext block has been modified, then that block will not be deciphered properly. In order to provide integrity, digital signature uses hash functions.

The secrecy of a public key is not required, but its authenticity is necessary to guarantee its integrity. If an attacker could replace the public key of an institution with his own public key, then the secret data intended for an institution would be accessible to the attacker. The attacker could receive and decrypt the message, and then send it on to the institution. Attacker could, also intercept the communications of two parties and make them believe that they communicate to each other, although they are talking to the intruder. A certificate authority can sign the public key of each user. When the network is established, when a user is added to the network, or when a user changes the user's public key, the public key is certified by a certificate authority 11. After certification, the user can send the certified public-key to another users who can then check its authenticity.

In order to become communication more secure, confidential and reliable there is a strong need to protect data and information broadened by computer networks and Internet. E-mail service provides users to exchange messages but sending an e-mail message is like sending a postcard that anyone can read. When an e-mail message travels from one e-mail server to another, the e-mail is first stored in an e-mail server before it is sent to the next e-mail server. A way to protect e-mail is by using writer-to-reader security in which the message is encrypted at the sender station and deciphered at the receiver station.

3. Conclusion and further work

Cryptography is the science of information and communication security. It is used for authentication and encryption (bank cards, e-commerce, smart health care cards, wireless telephone, pay-TV, etc.), access control, payment and it is becoming the basic tool of democracy with introduction of e-voting systems. In the last few years, a novel digital currency and peer-to-peer payment system named as Bitcoin is emerged. It uses cryptography to control, create and transfer of money and thus it is called cryptocurrency. Users send payments by broadcasting digitally signed messages through the network.

In the last 3-4 decades numerous modern encryption algorithms and systems are developed. It is accompanied by development of cryptology techniques. It means that every novel cryptosystem is subject of various attacks making the cryptosystems to be vulnerable. As a result, several novel scientific sub-disciplines appeared such as: quantum cryptography, post-quantum cryptography, DNA cryptography, cryptographic and secure engineering. Future works of development of new cryptographic algorithms should be in above-

mentioned sub-disciplines to make cryptosystems to be more resistant of cryptanalytic attacks.

References

1. P. Domosi, *A Novel Cryptosystem Based on Finite Automata Without Outputs*, *Automata*, Formal Languages and Algebraic Systems, AFLAS 2008, Kyoto, Japan 2008.
2. F. Seredynski, P. Bouvry, A. Y. Zomaya, *Cellular automata computation and secret key cryptography*, Parallel Computing, Vol. 30, Issues 5–6, 2004, pp 753–766.
3. R. Tao, *Finite Automata and Application to Cryptography*, Springer 2009.
4. B. Kar, D. C. Rao, A. K. Rath, *Generating PNS for Secret Key Cryptography Using Cellular Automaton*, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 5, 2011.
5. E. Franti, M. Dascalu, *Hardware Implementation on FPGA Platform for Cellular Automata Cryptosystem*, 5th WSEAS International Conference on Telecommunications and Informatics, Istanbul, Turkey, 2006, pp 330-333.
6. C. S. Rao, S. R. Attada, M. J. Rao, K. N. Rao, *Implementation of Object Oriented Encryption System Using Layered Cellular Automata*, International Journal of Engineering Science and Technology (IJEST), Vol. 3 No. 7, 2011.
7. Z. Feng, Y. Zhang, Y. Liu, *Image Encryption Based on Reversible Cellular Automata*, Advances in Informaton and Systems Sciences 2007 Vol. 2, pp 178-186.
8. I. Amorim, A. Machiavelo, R. Reis, *On Linear Finite Automata and Cryptography*, Technical Report, Universidade do Porto, PORTUGAL, 2011.
9. M. Seredynski, K. Pienkosz, P Bouvry, *Reversible Cellular Automata Based Encryption*, LNCS 3222, 2004, pp. 411-418.
10. F. Sereynski, P. Bouvry, A. Y. Zomaya, *Secret Key Cryptography with Cellular Automata*, Parallel and Distributed Processing Symposium, 2003.
11. Serge Vaudenay, *A Classical Introduction to Cryptography*, Application for Communication Security, Springer Science+Business Media, Inc., 2006.

Learning Management System (LMS) Usage: The Acceptance from Academicians Perspective

Saraswathy Thurairaj[1] , Komathi Munusamy[2] , Farida Bhanu bt Mohamed Yousoof[3]

[1] Universiti Tunku Abdul Rahman
tsaraswathy@utar.edu.my

[2] Universiti Tunku Abdul Rahman
komathy@utar.edu.my

[3] Universiti Tunku Abdul Rahman
faridabhanu@utar.edu.m

Faculty of Creative Industries & Faculty of Accountancy & Management (FAM)
Lot PT 21144, Jln Sungai Long Bandar Sg Long, Cheras 43000 Kajang Selangor

Abstract. The development of Information and Communication Technology (ICT) and in order to compete globally, learning and teaching landscape have changed tremendously. Driven by the explosive need for change and to produce excellent human capitals, Institution of Higher Learning are implementing changes by introducing technology as part of their teaching and learning process. Hence Learning Management System (LMS) has become essential part of small and large universities and able to provide series of capabilities and features. LMS permits users to access and deliver training content, leveraging a variety of training media throughout the organization and enhance the knowledge. The objective of this study is to examine whether academicians are ready to maximize the utilization of such technologies. This will provide insights on how Institution of Higher Learning can plan ways to promote readiness and acceptance of technology in teaching

Keywords: Learning Management System, Innovation in Education, Institution of Higher Learning

1. Introduction

Organizations recognize that the key element for survival in change. A change in organization is needed for continuous improvement simply because; the competitive global economy poses incessant threats to organization. Thus, organizations force themselves to constantly stay with the growing competition, globalization of markets, technology change, or even changes in government regulations by changing and innovating internally. The kind of change that occurs internally may involve adopting new technologies, redesigning business processes, new product development, changing employees' perception and thoughts as well as restructuring management structure (Ludon & Laudon, 2004). In fact, there may be other reasons that force or induce change, however, whatever the reason is for change to take place; it can't be denied that "change" do often have impact on the employees.

Although the need for change is evident that organization are trying to improve its performance through change process, yet employees' support or acceptance for change initiatives is totally hard to achieve. For instance, previous research has widely acknowledged that many organizational change initiatives neither results in their intended aims nor foster sustained changed (Choi, 2011). The failure is partly because the individuals within the organization do not support the change element. Thus, the change agent (leaders

and managers) should realize that it is important to assess not only the readiness level but also the acceptance of their employees towards change. Furthermore, individual differences among employees within the organization cannot be ignored as it is an important antecedence and they should be examined. These individuals should in fact made understand why change is needed, thus, making them ready to change their status quo to successfully accept the change.

Coming back to resistance to change, resistance to change is a common term used on organization change research. At all time, organization creates ideas to successfully implement change either by introducing new technologies, change of production processes, redesigning business processes and restructuring management (Julta, Craig & Bodoril, 2001; Ludon & Laudon, 2004), and new product development. All these are some change efforts to improve competitiveness by remaining competitive in business environment. Organizations therefore realize that change is necessary in today's rapidly changing business environment, however, majority of the change efforts fail (Beer & Nohria, 2000). As a matter of fact, these failures can be directly linked with resistance to change.

Additionally, Conner (1998), stressed that organizational change failure is also due to lack of attention to the "human side of change". This is said so because, organizational change initiatives are less likely to be accomplished when organizations fail to properly manage employee's reactions to the change process (Piderit, 2000). These individuals are an important element and in fact , a number of research have claimed that organizations only change and act all the way through their members and successful change will persist over the long term only when individuals modify their behaviors in a proper way (Jones, Jimmleson, & Griffiths, 2005; Meyer, Srinivas, Lal, & Topolnysky, 2007).

Considering the fact that individuals are an element in successfully implementing change, thus, it is ultimately important for organization to understand it. This is said so because, individuals' readiness and acceptance to change helps to sustain organizational efficiency, effectiveness, and technological improvement as well as a viable solution for long term organizational success and survival. Therefore focusing on these factors which help to support organizational change is necessary. Furthermore, each and every individual experience changes in a unique way (Bouckenooghe, 2010) and therefore, the variety in perception and reaction among employees within the organization cannot be denied. In brief, understanding the human side of change to resistance can increase the chances for organizations to be successful in their change initiates.

Basically all sectors within the economy are subject to an increasing need for change. For instance, Wittenstein (2008) highlighted that the health care industry need to keep up with the changing environment by adapting new medical technologies to cope with the increase demand in patient, shortages of doctors and many more. These scenario is no exceptional to the education industry specifically the institutions of higher learning (IHLs). IHLs are also in the stage to embrace change. The main objective of these educational institutions is to constantly strive towards quality teaching in order to produce excellent human capitals that are required by the country (Arokiasamy, Ismail, Ahmad &Jamilah; 2009). For that reason, these educational institutions are challenged to produce quality human resources, as such these institutions need to constantly introduce change and successfully implement it. Aydogen, (2007), in his research had mentioned that the reasons why education organizations need to change at all time is because they too need to keep up with the developments in the external environment and to produce human resources that are capable to compete from every angle within the corporate world.

Driven by the explosive need for change and to produce excellent human capitals, IHL(s) are implementing change by introducing technology as part of their teaching and

learning process. For instance, Mouza, (2002), explained that technology being introduced at classrooms increases potential for interaction and media rich communication as well as powerful educational tools that capable of providing excellent service to students. Recognizing the importance of technology in teaching and learning, many IHLs had formulated policies that require each and every academician to enhance the usage of technology to support student's learning. However, the actual usage of such technology by academic staff is still questionable.

2. Problem within the Institution of Higher Learning

IHLs are making it compulsory for academicians to incorporate learning management systems (LMS) into their traditional teaching and learning. With that in mind, it could be valuable to these IHLs to begin to understand the construct that contributes towards readiness of their academicians to accept organizational change. This is because, change itself is a stressful experience for many employees and for that reason two third of these change projects fails (Beer & Nohria, 2000).

Although it has been accepted by researchers that individuals are an important element in a change process, yet it can't be denied that organization change efforts still fail due to the inability of the organization to provide an effective unfreezing process. In a simple word, employees must be ready to adopt and willingly accept such changes so that the implementation of leaning management system can be done smoothly and successfully.

Hence, organization leaders and decision makers must understand the individual academicians and how to develop readiness among them, which in return, will be used to develop not only readiness but also acceptance towards technology usage in their teaching methods. In fact, past researches have highlighted that change agents underestimated the role individuals play in the change process (Armenakis, Harris, Mossholder, 1993; George & Jones, 2001).

A review on the past literature clearly states that successful implementation of organizational change lies within the role of individuals. A number of past researches had actually supported the importance of this factor, and thus, focused their research on various attitudinal constructs towards change. Therefore, it is an undeniable fact that individual level of analysis is important. With this in mind, it could be valuable to IHLs to begin a detailed concrete analysis on acceptance of academics as a constructs. Furthermore, study on readiness were published primarily in the health, psychology, and medical literature, however not many had purposefully looked into the context of academicians and to see whether this category of employee has the capacity to accept change.

3. Innovation in Education

Technology had been introduced into public universities in Malaysia since 1996 (Puteh, 2002). This idea behind this technology revolution was initiated and introduced by the former Prime Minister of Malaysia, Tun Dr. Mahathir Mohamad, and was followed by the establishment of the Multimedia Super Corridor in the year 1996 (Puteh, 2002). Beginning from here, the Malaysia education sector has evolved by promoting new approach in the practice of teaching and learning. In other words, the digital technology provides wide opportunity for educators to integrate technology-supported materials into their teaching-learning process. Therefore, the question here is whether academicians are ready to maximize the utilization of such technologies.

Besides that, another restructuring that is taking place in the educational sector is to have a pool of qualified academicians. The reason behind this is because, many universities and colleges began to change their institutional practices relating to teaching, structural learning, curricular and co-curricular initiative and multi-level assessment, as such having a pool of qualified academicians is totally needed. Therefore, a significant organizational change can be seen in IHLs which directly impacts this category of human capital specifically in the area of their professionalism. With that, academician's requirement had changed merely from just teaching to be talented and skilled professionals in their own area of expertise. This requirement is in line with IHLs movement to measure the performance of the academic using the Key Performance Indicators (KPI).

KPIs have covered various aspects which include: Teaching and Learning, Student outcomes, Educational Support Resources, Research, Publication and Training, Community Service and Administrative engagements. Therefore, it's clear that one of the performance measures used in KPI is the usage of educational support resource which addresses the technology aspect. For instance academics are required to improve their teaching techniques by implementing ICT in their teaching procedures and creating innovative and creative teaching methods. Besides having or compulsory to enhance their career development by participating in conferences, involving heavily in publication and consultancy, contributing their expertise in the administration of the university, as well as excellent teaching, they are now required to equip themselves with IT tools. Only with all these involvements they are now able to achieve a rating score in order to fulfill their KPIs.

No doubt that the commencement of IT and the requirement of academicians to cope with these digitalized technologies are interconnected. With the advancement in information technology, academic staff are often required to utilize in the teaching process. Despite several initiatives such as training and development (Nanayakkara, 2007) which has been aimed at improving academic staff skills in adapting and utilizing the IT technology, the success of integrating the system is still questionable. As such , the extent to which academicians in IHLs are compiling with this new requirement despite having heavy teaching and research requirement calls for a comprehensive analysis to determine the acceptance of the academic staff to maximize the benefit of such technology, merely, the Learning Management system (LMS). It is undeniable fact that technology interactive into teaching-learning is a must; however, academic staff's reaction to change and utilization of this technology is still questionable.

A. Issues in Technology Acceptance

There has been a continuous pressure on academicians to utilize technology to enrich their teaching process. This is said so because, IHLs are making it compulsory for academicians to incorporate learning management systems (LMS) into their traditional teaching and learning process. This organizational change element had actually forced the academic staff to revamp their way of delivery mode in order to react to the needs of the management. However, a clear problem that can be seen is that, the acceptance and maximum utilization of such system by overcoming resistance to change (RCT). Although the management had put it clear that the adoption of such technology change is visible, success of such a system is still questionable. Knowing the fact that LMS has become the requirement for their KPI, however, many academic staff are still reluctant to maximize the benefit of such tool. Alias, Nor Aziah & Zainuddin (2005) clearly state that it's a common scenario where academic staff are not fully utilizing the instructional technology. As such this situation leads to a scenario whereby they reject technology and totally resist change.

The success of LMS can only be seen if the academic staff enhance their desire to use the given technology (Ahmed, Zbib, Arokiasamy, Ramayah & Chiun, 2006) to the maximum. In other words, the success of technology change highly depends on the academic staffs' attitude to willingly accept the change and fully utilize to its benefits, subsequently reduces resistance to change (RTC). LMS consist of advance technology tools from uploading teaching materials to interactive features for the means of enhancing teaching; however, these advance features such as chat, discussion, forum, emails are not being utilized. There are many reasons why such phenomenon occurs, thus there is a need to further analyze the factors why academic staffs are reluctant to utilize the change as its benefits are greater.

B. Learning Management System (LMS)

As mentioned above, LMS consist of advance technology tools from uploading teaching materials to interactive features that helps to enhance teaching and learning. There are many forms of LMS such as Blackboard, WebCT, Moodle or Sakai (Jones & Muldoon, 2007), and these forms of LMS have been perhaps the most widely used educational technologies within universities and colleges, apart from the Internet and common office software (West, Waddoups, & Graham, 2006).

Nevertheless, the quantity and quality of learning occurring within these systems remains limited. Experience from one Australian university shows that as late as the second half of 2006, after over six years of institutional use of LMS, only just over half of all courses offered had course websites (Jones & Muldoon, 2007). This findings show that academicians are reluctant and perhaps not too keen to divert themselves from the traditional teaching methodology.

However, on the other hand, today's generation are known as "digital native" as they have matured in an immersive computing environment. With that, education industry rethinks ways is innovative in helping their students to acquire knowledge and skills. The advent in technology had created new technology changes which significantly impacts IHLs, specifically on teaching methodologies. For instance, in developing learning resources that can be transmitted via ICT or the internet. This will enable IHLs to keep up with the new trend in the knowledge economy.

Therefore, the significant growth of information can generally explain as an enriched technology enabled learning. By the end of 2003, many IHLs have installed some form of computer network. It is the first step to move beyond the traditional pedagogical process of learning. With that, over the years, many IHLs have incorporated LMS to facilitate teaching and learning. Ever since then, LMS have shown an obvious shift in the aspect of promoting variety of information and communication technology into the learning environment.

Incorporating LMS into facilitating learning environment is not only must but also necessity in order to be competitive in today's competitive environment. This is said so because by incorporating LMS, IHLs are now able to compete in the rapidly growing development in information and communication technology (ICT). Furthermore, as quality is a top requirement, LMS technology is believed to be able to maximize in improving the quality of teaching and learning. Therefore, academicians now have to take the responsibility to make the LMS a success. As such their creativity and innovative ideas are highly demanded specially in the e-content development.

Additionally, in order to ensure the quality of the programs being offered, the Malaysian Qualifying Agency (MQA) have been appointed to make sure that the IHLs administration and courses being offered are within the acceptable policies and procedures in

ensuring the quality of the programs offered. With respect to that, Malaysia education had undergone major changes over the years. However, besides the above changes in redesigning the programs and administrative structures, one of the emerging changes that needed to be addressed in today's environment is the growth and usage of technology in numerous ways by IHLs. However, the new technologies being introduced to improve the efficiency and effectiveness of the overall management of an university have not meet it intended objectives.

4. Conclusion

Change efforts are being initiated by MOHE effects academicians in the IHLs, for instance, introduction of technology in teaching i.e. LMS. However, academicians are underutilizing this advancement in technology in their teaching and learning process. Even with today's 'digital native' the success of technology usage like LMS is still questionable. Therefore, it is relevant to investigate the technology change efforts in the education setting by investigating the "individual construct". This will provide insights how IHLs can plan ways to promote readiness and acceptance of technology in teaching, i.e. LMS. Thus, these IHLs are also ahead of the advancement in IT usage in teaching. With that, a research on technology change is necessary for IHLs as they too are in a forceful and competitive environment.

References

1. Ahmed, Zafar U , Zbib, Imad , Arokiasamy, Sawaridass, Ramayah, T & Chiun, L, M,. (2006). Resistance to change and ERP implementation success: The moderating role of change management initiatives, Asian Academic Management Journal ,11,2.1-17
2. Alias, Nor Aziah ,Zainuddin, Ahmad,M., (2005). Innovation for better teaching and learning: adopting the learning management system},2, Malaysian Online Journal of Instructional Technology 227—40
3. Arokiasamy, L., Ismail., M, , Ahmad, A., Othman, J., (2009) Background of Malaysian private institutions of higher learning and challenges faced by academics, The Journal of International Social Research. 2, 8.
4. Armenakis, A. A., Harris, S. g., & Mossholder, K. W. (1993). Creating readiness for organizational change. Human Relations, 46: 681-703.
5. Aydogan A. (2007) Analyzing Corporatization Of User Generated Content In The Case Of Turkey
6. Beer, M. & Nohria, N. (2000). "Cracking the code of change," Harvard Business Review, 78 (3): 133-141.
7. Bouckenooghe, D., (2010). Positioning change recipients' attitudes toward change in the organizational change literature, The Journal of Applied Behavioral Science, 46(6), 500-531
8. Choi,M., (2011)Employees' attitudes toward organizational change: A literature review, Human Resource Management, 50(4), 479-500
9. Conner, D.R. (1998), Leading at the Edge of Chaos: How to Create the Nimble Organization, John Wiley, New York, NY.
10. George, J. M., & Jones, G. R. (2001). Towards a process model of individual change in organizations. Human Relations, 54(4), 419-444
11. Jones, D., & Muldoon, N., (2007). The teleological reason why ICTs limit choice for university learners and learning. Providing choices for learners and learning. Proceedings ASCILITE Singapore , 450-459
12. Jones, R. A., Jimmieson, N. L., & Griffiths, A. (2005). The impact of organizational culture and reshaping capabilities on change implementation success: The mediating role of

readiness for change. Journal of Management Studies, 42, 361-386.

13. Julta.D.N, Craig, J.& Bodoril, P (2001). A Methodology for creating e-Business strategy. Proceeding in the 34 Hawaii International Conference on System Science
14. Laudon and Laudon J. (2004). Management information systems: managing the digital firm. Management information systems, 1, 103-105
15. Meyer J.P., Srinivas, ES, Lal, J., B, Topolnytsky, L (2007). Employee commitment and support for an organizational change: Test of the three-component model in two cultures}, Journal of Occupational and Organizational Psychology. 80(2), 185—211
16. Mouza, C., (2002), Learning with laptops: The impact of one-to-one computing on student attitudes and classroom perceptions .Proceedings of the 7th international conference on Learning sciences, 488—494.
17. Nanayakkara C (2007). A model of user acceptance of learning management systems: a study within tertiary institutions in New Zealand, The International Journal of Learning
18. Piderit, S, K (2000) Rethinking resistance and recognizing ambivalence: A multidimensional view of attitudes toward an organizational change, Academy of management review 25(4)
19. Puteh, M (2002)., Qualitative research approach towards factors associated with mathematics anxiety., The 3rd International Conference on Mathematics
20. West, R., Waddoups, G., & Graham, C. (2006). Understanding the experience of instructors as they adopt a course management system. *Educational Technology Research and Development*, *55*(1), 1-26.
21. Wittenstein R.D (2008) Factors Influencing Individual Readiness for Change in a Health Care Environment.

Constructing Tests and Dialog Modeling Using Ontological Engineering.

Verlan A.F.[1], Furtat Yu.O.[2]

[1]National Academy of Sciences of Ukraine Pukhov Institute for Modeling in Energy Engineering, General Naumov Str., 15, 03164, Kyiv, Ukraine
a.f.verlan@gmail.com

[2] National Academy of Sciences of Ukraine Pukhov Institute for Modeling in Energy Engineering, General Naumov Str., 15, 03164, Kyiv, Ukraine
saodhar@ukr.net

Abstract. The paper is dedicated to the problem of tests construction and domain knowledge representation in automated testing systems. Problems to solve are named and erotematic dialog modeling method is proposed as a possible solution. For different types of tests corresponding classes of questions are proposed. Information technology of the user's knowledge level assessment is described

Keywords. Automated training system, knowledge representation, declarative knowledge, "question-answer" situation, derivation engine.

1. Introduction

During training (in education or in professional training) control of knowledge is carried out in periods set by the training regulations. Automated training systems (ATS) usage provides prompt control of knowledge, in theory – at any given time. One way to control the state of knowledge is testing.

2. Computer testing systems features.

An electronic version of the test kits, software tools for dialogue "tutor - student (user)" are required for computer testing as well as tools to support test sets (test database). Generally, computer-aided learning systems have developed services to enter and edit the tests' text forms. These tests are provided by teachers of the appropriate academic subject. They are the authors of manuals or books that contain some declarative knowledge on the subject required for students. If the teacher is not the author of texts that contain declarative knowledge, he refers to the appropriate manual or other sources. These sources are used to generate test cases for knowledge control.

ATS should include knowledge of the subject area. Obviously, to answer the question of whether the training goals are achieved one has to to compare the user's state of knowledge to the domain description. This description is also the base of comparison, and a source of training sequences.

Since the description of the domain is contained in ATS, it can be used to build a set of tests. To generate test sequences one has to solve the following problems:

- choose a model of the "tutor-student" communication;
- choose a model of exchanging elementary messages [1];

- support the student's knowledge model to determine which part of the training course has been learnt and can be tested;
- choose a form of domain knowledge representation;
- create software modules for dialog messages elements construction;
- create a domain-independent answer control mechanism.

3. Formal models of communication.

Psychologists consider the need for information one of the basic human needs, stemming from the need to reduce the level of uncertainty of his informational state. In the process of communication the individual uses perception mechanism to process the partner's notifications. The question is, as a mean of communication, often a linguistic form. The exchange of questions and answers generates a sequence of "question-answer" situations, forming a dialog. Analysis of dialog as a form of verbal communication is conducted from the time of Plato. Dialog will be considered a sequence of steps, each of which consists of a loop formed by two interrelated linguistic acts of partners. Agreeably, a dialogue consisting of erotematic (interrogative) situations will be called an erotematic dialog. For erotematic "question - answer" situations a special formal apparatus is developed – "erotematic logic".

If testing can generally be implemented as a dialog, and basic test can be viewed as an erotematic situation, one can put a formal framework for the implementation of computer methods for sets of test construction and response control. There are descriptions of the methods of constructing test for when they are formulated by a human teacher manually or by using test editor or tests database management system. For example, the requirements for the tests are proposed in [2]. A formal method for the construction of an elementary test would allow its use for a wide range of problem areas.

Let us show that testing for knowledge level control can be described as a "question - answer" situation . To do this, we assume that the "question - answer" situation corresponds to the term "test step". By this term we mean the test answers transfer in the form of messages to the listener and receipt and control of answers (only for compliance to the interrogative). In [2], the types of tests are offered and corresponding to these tests types of questions [3].

4. Comparison of question classes and test types.

There is a difference between questions and test tasks. Question incorporates for the one who answers the question all the information that is already known to the person who asks it. The question's goal is to get the missing knowledge. The purpose of the test task is different: the tutor aims to obtain knowledge about the state of knowledge of the listener.

To do this tutor adds to the list of alternatives for the answer a false alternative, known as "distractor". From the standpoint of erotematic logic, all alternatives contained in the question's subject are intended for designing an answer and can be transmitted to the asking partner, that is to construct a response. Distractor is used to model the student's knowledge perturbations.

Table 1
Matching question classes and test types

№	Test task type (by Kline)	Equivalent question class	Subject representation form	Question preconditions requirements
1.	Multiple choices task	"what kind of"-question	explicit list of alternatives that has no more than 5 items	it is required to choose no more than one alternative
2.	Alternatives task	"either..or"-question	two alternatives: "yes and no"	the same
3.	Bezel question	reduced to the "what kind of"-question	silenced implicit list of alternatives	the same
4.	Tasks with template answers	reduced to the "what kind of"-question	implicit by default list of alternatives	no more than one alternative with a given number of characters
5.	Task with a set of keywords (images, symbols) from which to construct an answer	can represent “what”, “where”, “when”-questions that are reduced to the "what kind of"-question	explicit list of alternatives	no more than one alternative
6.	Closed form questions	the same	explicit list of alternatives	no more than one alternative, the answer is given as a character, that names an alternative (no icon)
7.	Compliance task	reduced to the conjunction of “either..or”-questions and their subjects	explicit list of alternatives	no more than one alternative
8.	Constructing a reply	reduced to the “either..or”-question		
9.	The demonstration task with moving objects	reduced to the “either..or”-question	subject represented by scales	no more than one alternative

5. Declarative knowledge representation.

Knowledge representing models designed to determine the success of training described in [4, 5].

Ontology represents the subject area as a combination of a hierarchy of objects-classes, properties of abstract objects and relations between them. In [4;6-8] the proposed method of knowledge representation and questions construction for the user based on this scheme.

To simulate the student's knowledge, "information phantom" dynamically displays the appending of each student's knowledge for each domain category, intelligent tutor software components monitor the results of testing the success of knowledge mastering. Depending on the results of testing the level of uncertainty of student's knowledge elements model changes. Another test question is formed by the tutor software component taking into account the student's knowledge element with the lowest degree of certainty. Thus is formed a link between the current and the next dialog "question-answer" situation or, in other words, between the current and the next test question.

The construction of the test in the form of "question-statement" requires a mechanism of logical reasoning (derivation engine in the context of building intelligent systems). For a derivation machine a knowledge representation model is required, for which such a machine is possible. This model should reflect knowledge about the subject (domain) and the state of knowledge of the listener. The purpose of education can be interpreted as monitoring students' knowledge so that it is, ideally, a copy of the domain knowledge model.

6. Forming of the domain knowledge model using ontological engineering methods

Ontology provides a dictionary of terms and concepts to describe the subject area and a set of boolean expressions that define constraints and interpret dictionary.

To construct the ontological description of the domain while taking into account the standard IDEF5, one must perform such procedures [9, p. 22]:

- use of existing ontologies;
- create (edit) glossary of terms;
- definition of basic concepts and relationships;
- selection or creation of a special language for describing ontologies;
- formation of hierarchies of classes (taxonomy).

In [10] the concepts are defined as means of identifying the knowledge, they have two main functions – to recognize and to generate models of the elements of the world (11, p. 23). Models generation is the basis of the objects design (engineering). Knowledge acquisition means classification contains a method of "formation of concepts".

To form a concept, the following operations are used: comparison, analysis, significant parts accentuation and synthesis. In order to define some concepts it is needed to name a general concept (family, supersystem), which contains this concept, and to find such qualities (properties, essential features) that are common to all objects, named by this this concept, and do not belong to the other concepts of the supersystem.

The complexity of the task is associated primarily with the concepts set great cardinality and ambiguity. Erotematic approach makes it possible to solve this problem in formulating the questions that would guarantee getting correct answers regardless of the question variables lexical meanings. False answer values should also be determined regardless of their values. This will provide an opportunity to implement dialogical erotematic machine that can work with the "question-answer" situations.

Thus, information technology of the user's knowledge (professional or acquired during the educational process) level assessment consists of the following steps.

1.Domain ontological description formation, which consist of a network of concepts and corresponds to the domain declarative knowledge. Forming is performed by the tutor (or expert) using adaptive information system that implements a some meta-script.

2.For each user declarative knowledge display is supported, which corresponds to a network of concepts, and includes: a) the entropy values for object-property-relation categories; b) the user's cognitive type description.

3.Test questions are formed by the DM 2 unit (interactive machine of the second kind) using standard relations.

4.The answer to the test question is processed by the DM2 as part of the standard "question-answer" situation (using disreactors).

7. Conclusions

The proposed approach allows to create a system of testing and assessing the level of knowledge for any domain for use in automated training systems. If there is a formal description of domain ontology tests creation process can be further automated.

References

1. Chimir I. Dialogue of Partners as a Method of Non-Formal Problem Solving / I. Chmyr ; In eds. D. Maddy, J. Brouwer and L. Thomas Harrington // Human-Machine Communication for Education System Design – 1994. – № F129. – P. 221–228. (In English)
2. Kline P. Introduction to psychometric design : [reference manual for the construction of test]. – Kyiv: Pan Ltd., 1994. – 284 p. (In Russian)
3. Belnap N., Stil T. Logic of questions and answers. – Moscow: Progress, 1981. – 288 p. (In Russian)
4. Ous M.F., Gadezka Z.M. Knowledge representation model for student testing // System analysis, management and information technology. Bulletin of the Kharkov State Polytechnic University : [collection of scientific papers]. – Charkiv, 2000. – Iss. 94. – P. 82–86. (In Russian)
5. Ous M.F., Kostian N.L., Tarandushka L.A. Using CASE-technologies for domain knowledge modeling in distance learning // Means and methods of "e-learning" in continuous education : interuniversity scientific-method. conf., 26-27th of April 2006. : proceedings abstracts. – Cherkassy, 2006. – P. 29–37. (In Russian)
6. Ous M.F., Piskun A.V.,Gadezka Z.M. Sudent's cognitive level modeling in intelligent tutoring systems // Modeling and informational technologies : [NAS of Ukraine Pukhov Institute for Modeling in Energy Engineering scient. works digest]. – Kyiv, 1999. – Iss. 3. – P. 133–141. (In Russian)
7. Ous M.F., Gadezka Z.M., Kostian N.L. Methods of infological and ontological modeling for representing knowledge in training // Modern management in manufacturing and humanitarian activities : Intern. sciences. conf., April 2007 p. : proceedings abstracts. – Cherkassy, 2007. – P. 35–36. (In Russian)
8. Ous M.F., Gadezka Z.M. Evaluation of student's cognitive level in personalized computer-based training systems // Information Technologies in Science, Education, Engineering" (ITSEE-2002) : All-Ukrainian. conf. of young scientists, 17-19th of April 2002. : proceedings abstracts. – Cherkassy, 2002. – P. 157–160. (In Russian)
9. Palagin A.V., Yakovlev Yu.S. Building a domain ontology " Intelligent informational systems " // Control Systems and Machines. – 2005. – № 6. – P. 18–27. (In Russian)
10. Gladun V.P. Partnership with PC. – Kyiv: Port-Royal, 2000. – 128 p. (In Russian)

Some Properties of Processes Acoustic Emission Control Structures

M.M. Karimov, M.V. Sagatov, B.A. Hudayarov, R.M. Irmukhamedova

Tashkent State Technology University, Uzbekistan,
informtgtu@mail.ru

Abstract. It is shown that now nondestructive control of materials and products a measuring method of acoustic emission is directly connected with a problem of recognition of stages and mechanisms of education and growth of defects in samples. The frequency range of signals acoustic emission contains important information about proceeding in studied object. In work change of a range of frequencies of signals acoustic emission at various stages of development cracks in a sample and its communication with coefficient of intensity of tension is studied. It is established that in the course of growth of cracks there is a narrowing of width of a range frequencies of signals acoustic emission. The received theoretical results well correlate with data of the pilot studies executed on a complex of blocks measurement of signals acoustic emission which allows automate process of collecting, processing and registration of measuring information.

Keywords: acoustic emission, the spectral analysis, characteristics of signals acoustic emission, transformation process in mechanisms and their knots, plastic deformation, education and growth of defects in samples.

1. Introduction

Non-destructive testing of materials and products using acoustic emission (AE) in its development into a dense approached the problem of the recognition of the stages and mechanisms of formation and growth of defects on the AE signals. For this purpose, since the 70-s, was developed spectral analysis of the AE signals.

Some theoretical calculations of the spectra of AE signals arising from the destruction of the materials showed that the range of these changes can be up to several megahertz [1-4]. Practical use of such a broad band spectrum of AE signals is difficult because of the complexity of its implementation in the measurement technique. However, in this and there is a great need. As the experimental results for these purposes it is sufficient information is the frequency range up to 4 MHz [5, 6]. In [7], where it was used technically well-equipped measuring base and a wide range of various sensors indicated that:

1) primary energy spectrum of AE signals is concentrated in the range below 1 MHz, and in many materials , a sharp peak appears AE signals in the range of 100-150 kHz;

2) the highest peak of the frequency spectrum (with a maximum of AE- radiation) was recorded at a frequency of 350 kHz . The paper also noted that the experimental results do not depend on the type of material (steel, samples of HF- 1 and brass brand 70-30) , AE transducer , sample geometry and harmonic analysis method. However, in a number of papers claimed that the spectrum of AE signals contain not only information about the physical processes that cause AE, but also that the shape and dimensions of the sample, and the characteristics of piezoelectric transducers significantly alter the original signal AE [3].

From these results it is evident that the frequency spectrum of AE signals contain important information about running processes in the controlled object, but in some cases the

data are contradictory. In addition, the literature is poorly lit question regard the spectral characteristics of AE signals with the stages of development of cracks and its mechanisms. These issues are discussed in this paper.

2. Theoretical background

Establish a change of the frequency spectrum of AE signals at different stages of crack development and its relationship to the stress intensity factor. As is known, an electric voltage U (t) at the output - the recording apparatus AE can be represented as the convolution of

$$U(t) = \int_{t_0}^{t} u(\tau)v(\tau - t)dx, \quad (1)$$

where, *u (t)* - integral mechanical movement of the transducer under the AE; *v (t)* - the transfer function of the receive path of AE equipment.

If you apply to (1) the integral Fourier transform and the convolution theorem, the relationship between the spectral characteristics can be summarized as follows:

$$\overline{U}(f) = \overline{u}(f)\overline{v}(f), \quad (2)$$

where the dash on top of the functions is the Fourier transform of these functions ; *f* - frequency.

As can be seen from the relation (2) at a known transfer function can find the value or vice versa, in a certain frequency range.

In practice, it often happens that the function is unknown. We show that even in this case the relation (2) we can obtain important practical results. Indeed, if the position of the transducer on the subject of research is not changed and collect information from the same control zone (which can be determined using triangulation techniques), then the value will depend on the stage and the mechanism of the defect.

Consider the most dangerous in terms of the strength of products defect such as a crack. In the case of formation of the inner disk-shaped cracks form of elastic waves (the function u (t)) and the modulus of the frequency spectrum (function) are given in [1, 2]. The main emission peak in the far-field (maximum) at 0.7 of the maximum amplitude value of a width for metals that Poisson's ratio close to 0.3. Assume that a first approximation for waveform formation of cracks and crack jump to the same value as the newly formed, identical, differing only in amplitude. Then the width of the frequency spectrum with a single jump of the crack

$$\Delta f = 0{,}125\alpha c_2 / \pi r_0, \quad (3)$$

where a - proportionality factor between the width of the frequency spectrum in the formation of cracks and jump; c_2 - shear wave velocity; r_0 - radius of the disk, equal to twice the area of the newly formed as a result of the jump surface; f - frequency.

According to the theory of a single averaged value of the jump can be written as

$$\Delta l_{cp} = r_0 = \alpha K l_{cp0}, \quad (4)$$

where α - constant of the material - the average value of stress intensity factor.

Then jump in the crack area (the area of the semicircle)

$$\Delta s = 0{,}5\alpha^2 K l_{cp}. \quad (5)$$

Substituting (4) into (3) we have

$$\Delta f = \beta K l_{cp0}, \tag{6}$$

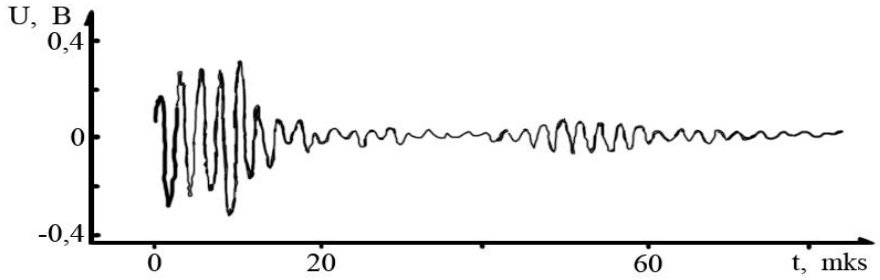
where β - 0.25 ac2 / a, ie the width of the frequency interval back lroportsionalna square of the stress intensity factor, all other things being equal.

As the experimental studies , the value of (at Breakaway cracks) and (maximum stress intensity factor} relate to each other in the range of 1:2.3 to a variety of materials, ie, an average of at 1:1.9. Therefore, according to (6), the width of the spectrum frequencies beginning to crack growth in the critical transition stage of its development will be on average 3.6 times smaller. example, relates to a titanium alloy as 1:1.35, i.e. in the final stage of crack width spectrum decreased by (6) of 1.82 times compared to its start.

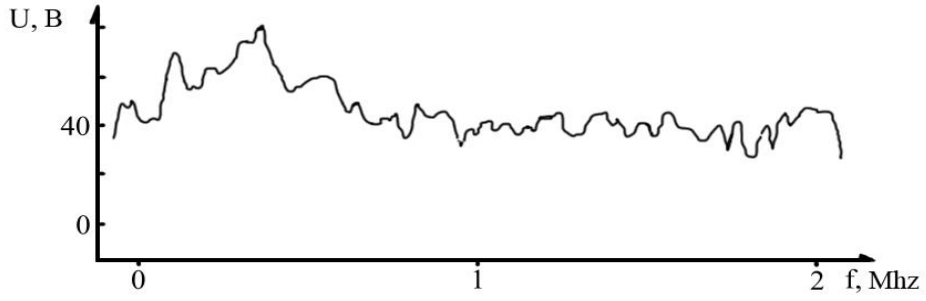
The width of the frequency spectrum at the value according to (6) is

$$\Delta f_c = \beta K_c^{-2}. \tag{7}$$

From the theoretical results that during the crack growth is narrowed spectrum width of the AE signals. If, during the development of the crack measure this parameter, and compare it with the value calculated by the formula (6), it is possible to predict the destruction of material or structural element. For example, when the operation of the structural element is terminated. Therefore, the width of the spectrum can be used as the parameter information for the diagnosis of operating materials and structural elements. We show that the theoretical results are in good agreement with the experimental data.



a)



b)

Fig. I. Calibration signal from the fracture of a glass capillary (a) and derived from this frequency response AE transducer (b)

For experimental studies used serial piezoconverter elastic waves into electrical signals from the kit AE-15 unit AF preamplifier, complex measurement units AE signals. Software made in complex software environment Matlab.

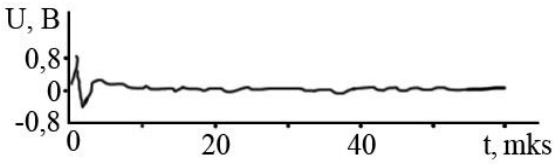
AE calibration signal and the amplitude-frequency response of the converter shown in hig. 1. From fig.1 shows that the highly satisfactory to evaluate the uniformity of the spectral characteristics of AE signals is the frequency band 176-500 kHz , which is consistent with the recommendations of [8] for the selection of the optimal range of the frequency spectrum of AE signals . However, the converter allows you to receive and AE signals of higher frequencies up to several megahertz.

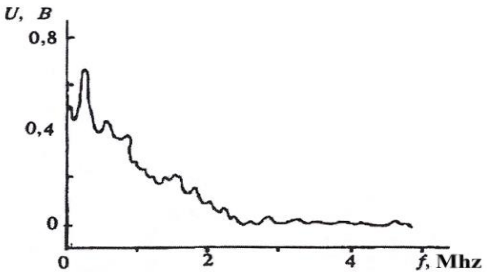
QRS complex to study the spectral characteristics of AE signals will automate the process of selection, processing and recording the results. It also provides timely information about the basic parameters of the spectra of AE both visually and by printing them on a printer.

3. Measurement of amplitude-frequency characteristics (AFC) system sample converter

Based on the experimental data and their analysis shows that the acoustic field generated by a point source locally homogeneous, and the frequency response of the system to study converter has the properties of ergodicity. Therefore it is recommended to obtain the frequency response of the system to initiate the calibration point source AE- random process.

We succeeded in the process of calibrating the system inverter -sample - loading device to obtain short pulse signal AE during loading of the sample. His appearance and spectral characteristics of this system are shown in fig.2. Such a signal was recorded from a single short jump crack. The signal is of short duration, so the maximum frequency response shows the spectral properties of the system. Pulse duration of 3.5ms., and the scope of his output measuring complex was equal 1.33 V.





a) b)

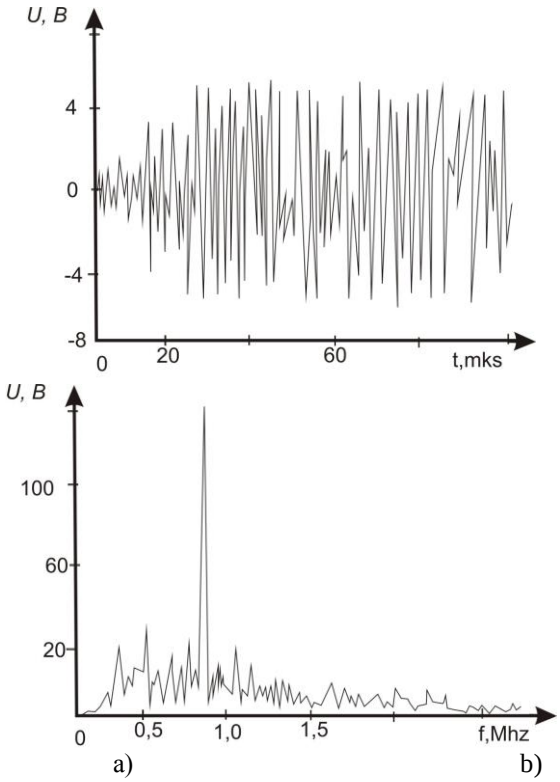
Fig.2. Calibration pulse of the jump of the crack (a) and the frequency response of the sample-converter-loading device (b)

As shown in Figure 4 the frequency response of the system, the most efficient registration process AE signals will pass frequencies up to 980 kHz, but the entire band frequency response lies within 0,01-1,9 MHz, which is sufficient for experimental studies.

4. The results of measurements and their interpretation

Spectral analysis of AE signals were carried out at various points in the diagram "load - disclosure of the crack." In the linear region of this diagram is the very first AE signals were recorded at low load P (respectively, the stress intensity factor). Such an AE signal and its spectrum is shown in fig 3.

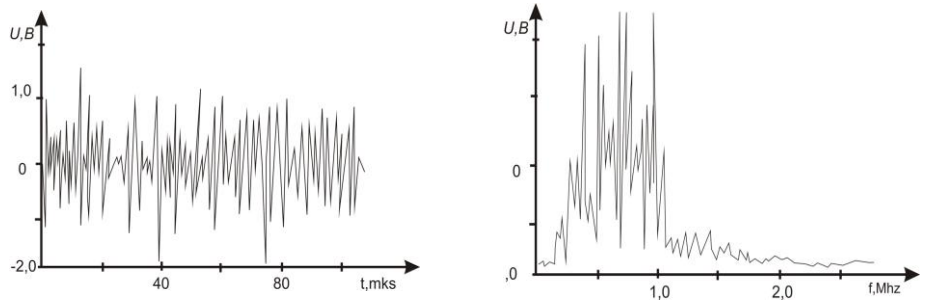
The spectral characteristics of the signal differ quite a broad band of approximately 2 MHz, and the peak emission occurs at a frequency of 0.62 MHz. Apparently, this signal corresponds to the growth of the plastic zone at the top of preliminarily derived fatigue crack and microcracking.



a) b)

Fig.3. AE signal obtained in the linear part of the diagram "load-disclosure of the crack" (a) and its response (b)

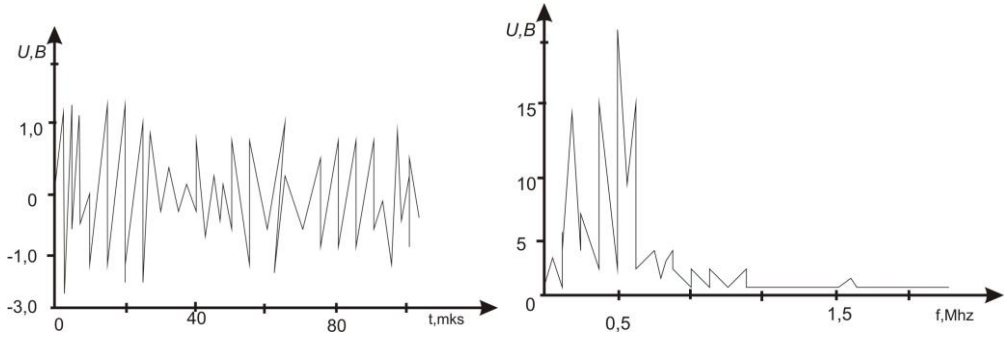
As the starting subcritical crack growth, which leads to a change in the spectral characteristics of the AE signals (Fig.4). There is a tendency to narrow spectral band response (0,14-0,96 MHz), as well as offset the emission peak in the low-frequency region.



a) b)

Fig. 4. Type of AE signal in the middle stages of subcritical crack growth (a) and his response (b).

In the neighborhood of the deviation from linearity in the diagram "load-disclosure of crack" observed AE signals with narrow band spectral characteristics and with a clear shift of the peak emission in the low frequency region (fig.5). These AE signals corresponding jumps microcracks. Upper band frequency response are becoming 0,09-0,53 MHz emission peak at a frequency of 0.37 MHz.



a) b)

Fig. 5. Type AE signal at the final stage of subcritical growth fracture (a) and response (b)

At the stage of subcritical crack growth obtained a good agreement between experimental and theoretical data. Figure 6 shows the curve changes the width of the frequency spectrum at different stages of the growth, built on the theoretical results (solid

line) and the experimental data (points) for the value β = 878,6 MPa2ms-1. This confirms the fact that the development of subcritical crack growth width of the frequency spectrum is inversely proportional to the stress intensity factor.

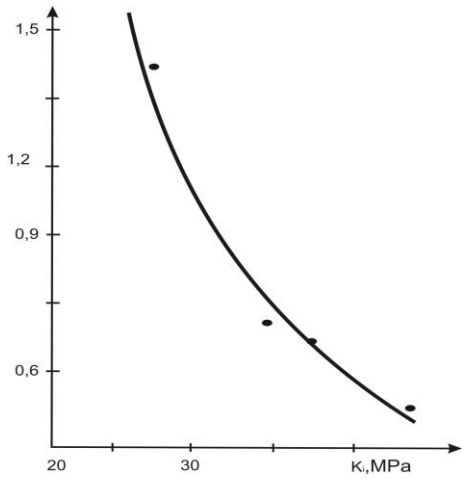
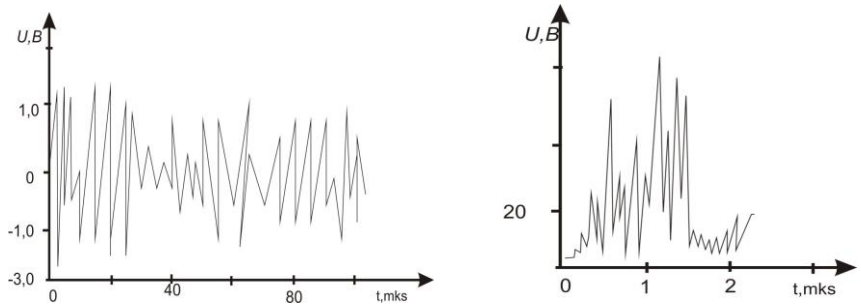


Fig. 6. Dependence of the width of the frequency spectrum of AE signals from the stress intensity factor

When passing through the point Pm (start supercritical stage of crack growth) is observed again expanding band frequency response and shift of the peak emission at higher frequencies (fig.7). This corresponds to a deviation from the straight path of crack growth, ie, in this case, we have replaced the failure mechanism of normal separation to mixed : normal takeoff - shift. This does not exclude the option that takes place when they are shifted friction of the crack. Macrodestruction development until the complete destruction of the sample is accompanied by wide frequency response of AE signals. Emission peaks fall on both low and high frequencies on. The width of the frequency spectrum was the same order as that at the initial stages of the normal separation Between the AE signals wide range of AE signals and met with a rather narrow spectra, which are characteristic of subcritical crack growth mechanism of normal detachment, suggesting that there is a crack jumps on this mechanism for supercritical stage of development.



a) b)

Fig.7. Typical AE signal (a) and its response (b) in step supercritical crack growth

5. Conclusion

The transition from the stage of destruction of plastic deformation and microfracture to the stage of subcritical crack growth is characterized by narrowing of the spectrum of AE signals.

The main part of the energy signal AE crack growth under normal separation in concentrated spectral range below 1 MHz. In this subcritical crack growth in the initial stages is different from its end stages of a further narrowing of the frequency spectrum, the width of which is inversely proportional to the stress intensity factor.

Fracture mechanism transition from normal separation to mixed peel- shift leads to a broadening of the frequency spectrum, which complicates the identification of this parameter stages of subcritical crack growth and the mechanism of its development.

References

1. Andreykiv A. E., Lisak N.V. Method of acoustic emission in the study of fracture processes. - Kiev: Naukova Dumka, 1989. – 176p.
2. A.E. Andreykiv, N.V. Lisak, O. Sergienko, V.R. Skalsky. Theoretical concepts of acoustic emission method in the study of fracture processes. / - Lviv / PMI them. G.V. Karpenko Ukrainian Academy of Sciences, 1987. - 49. (Preprint number 137).
3. Greshnikov V.A., Braginsky AP An analysis of acoustic emission signals / / Flaw. 1980. - № 5. - P. 101-106.
4. Verlan A.F., Goroshko I.O., Karpenko E.J., Korolev V.Y., Mosentsova L.V. Methods and algorithms for signal and image restoration / / KNAS Ukraine, 2001-369p.
5. Acquaviva C. J. Interlaboratory comparisaua of acoustic emission spec //NDT International. 1980. - № 5. - P. 230-234.
6. Muravin G.B., Lezvinskaya L.M. Study the spectral density of the signal AE / / Flaw. 1982. - № 7. - P. 10-15.
7. Trukhachev A.A. Analysis procedures and algorithms for detecting signals. M:. Radio and Communications, 2003. – 248p.
8. Gorbunov A., Lykov Y.I. Influence of amplitude-frequency characteristics of the object on the spectral characteristics of AE signals / / Flaw. 1986. - № 9. - P. 39-45.

About One of the Authentication Methods

Madjit Malikovich Karimov, Abduaziz Abdugaffarovich Abdurakhmanov, Nurbek Bakhtiyorovich Nasrullaev

Tashkent University of Information Technology, Uzbekistan,
dr.mmkarimov@rambler.ru, k.bakhtiyarovich@gmail.com

Abstract. This article proposes the issues of information security in Web-portals with the use of authentication technologies access to the remote user and an algorithm for authentication in an interconnected mode. Shows a comparative analysis of the proposed method with others.

Key words: Web-portal, authentication, threats, algorithm, protection.

1. Introduction

In accordance with modern principles of construction of information infrastructure interaction of state agencies or commercial structure with the citizens (customers), other organizations, including governments, it is advisable to ensure, through produced by the Internet news portal, which is a single point of access for all users, resources or, alternatively, to a set of information services.

The main objective of creating accessible web portals is to improve information support of the population of other consumers about the activities of federal and regional authorities, socio-economic development, legislative activities, etc.

A number of Internet portals that are created by state or commercial structures, the problem must be implemented to ensure free access for all users of the Internet to the shared resources, and organization of the authorized users of the portal with various services (applications). These problems solved by using portal technology, include the implementation of e-commerce systems, providing secure access for remote users to corporate resources, etc.

The problem of protection of information resources such distributed infocommunication system using open channels of communication (the Internet) can be solved on the basis of a comprehensive, integrated approach to information security of Internet portals. This is achieved by creating a single design process of infocommunication system protection of information resources (PIR) portal. [3, 4].

2. Problems of PIR protecting

Under PIR we mean a set of measures and tools aimed at the identification, reflection, and the elimination of security threats specific Internet portal.

The modern concept of object protection information, including Web portals, involves the implementation of the following sequence of actions:

- analysis of the functioning of information objects and resources (elements) that require the security of information (the definition of protected objects);
- identification of possible threats to their ranking in importance and the formation of a list of requirements to ensure the security of information protected resources;

- development of adequate information security threats measures, the choice of tools and technologies to protect;
- agreeing the choice of information technology (hardware and software, and network resources) in terms of applicability of information security and implementation based on PIR comprehensive approach to protect information resources Internet portal.

Please be aware that maintenance is PIR system to the portal. It must include adequate protection to all information technology (IT) to protect critical resources and information to the consumer without compromising quality of hardware and software info communication system.

This is achieved through a number of interrelated organizational and technical protection mechanisms to substantially reduce the probability of the main threats and create a platform for improving the system control and protection of information resources and capacity to secure Internet portal functions.

The information provided by users (readers) public Internet portals is open. Therefore the main task in implementing the security portal is to ensure the integrity and availability of their information resources. In this regard, by posing requires, first and foremost, protecting against unauthorized access from the Internet.

Thus, the most pressing problem is the implementation of strong authentication of authorized users on the portal for access to corporate resources, etc.

In particular, a key problem for information security is to ensure the OE integrity and reliability of circulating in the system for electronic documents. The presence of such a claim involves the inclusion of PIR portal software and hardware to ensure:

- provability of authorship made, modified and deleted records, generating electronic documents, including the placement of the portal content;
- authorized the introduction, modification and deletion of records in the portal administrators databases, etc.

The details of these requirements will be described below.

3. Structure of secure web portal

The structure of a typical secure Internet portal should be designed so that the technologies that have been implemented within the framework of poses, have been involved in all solved problems of information portal (portal administration, management of both public and internal, intended for authenticating user content; portal connection by means of telecommunications to the Internet, etc.).

The block diagram of a typical building PIR secure Internet portal is shown in next Figure 1.

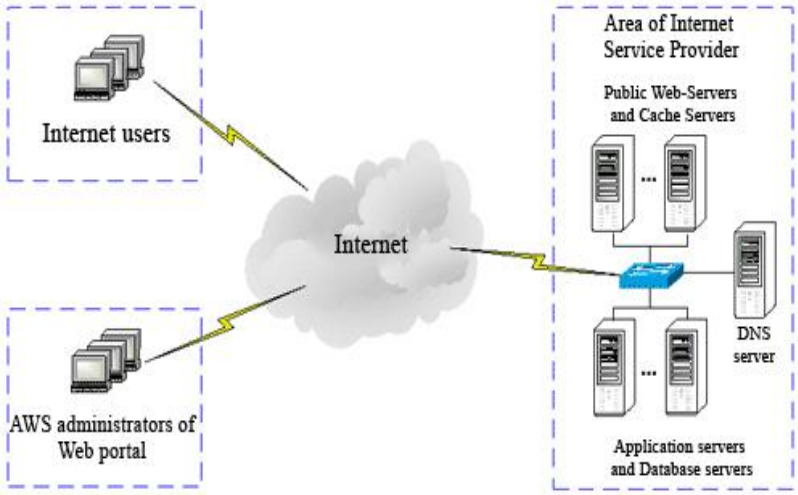


Fig.1. Typical architecture of Web-portal

In order to ensure comprehensive protection of information resources Internet portal model posing it should implement the following functions and include the corresponding subsystems:

- protection against unauthorized access;
- active audit;
- anti-virus protection;
- Cryptographic protection and support of digital signature (EDS).

The function of protection against unauthorized access is implemented using:

- protect against unauthorized access portal resources from the Internet and at the level of the protected LAN segments, including the regular means of protection used in the site operating systems (OS) and database management systems (DBMS). This includes the use of certified information security requirements OS and DBMS;
- Administration of protection against unauthorized access;
- System event recording and unauthorized access attempts to protected resources, early warning security administrators about unauthorized access attempts, etc.

4. Authentication process

In systems of authentication and access control is usually implemented as two entirely different things. Although sometimes it makes sense to distinguish between the bolt that holds the door closed and locked, which controls the valve, such things are often integrated into a single mechanism. In computers, the same authentication process simply sets up the correct user name for future use, and access control is carried out separately. Computer systems typically control access by comparing a user with access rules, tied to a particular file or other resource. If the rules allow access to the user with that name, then he gets the opportunity to use the resource. Current authentication systems have evolved as a result of decades of attacks, many of which were successful. This has led to attempts to intercept passwords, which, in turn, - to the emergence of other protective measures.

Algorithm of authentication. Mutual authentication of the user and the network, except for user authentication, and provides authenticated network. In the latter case the

object is authenticated network user. The figure shows a five-step scheme of mutual authentication system (B) and a guest user of the system (A) services to info communications systems.

T - mobile user. In the message-1, a system of info communications services system B sends a random number RB in the guest user to the system A. Guest system translates into a mobile user that the number of T, and transmits the generated number R (message-2).

From T mobile user in the guestbook will be reported by message-3 KTA(R, KS, KTB (RB, R, KS)). Message-3 is encrypted shared key KTA mobile user and guest systems UPT. This report includes:

• receipt KTB (RB, R, KS) access to the system B. Here KS - session key generated by the mobile system;

• a random number R;

• session key KS.

Message-3 stands for the guest user's system and the key KTB. The random number R authenticates the mobile system T. Guest user system transmits receipt system (message-4).

The system decrypts the received ticket, receives and compares with RB, sent in a message-1. The coincidence of these values confirms the authenticity of the guest user's system.

In the system sends a random number R, the encrypted key KS (KS key is decrypted when receiving a message-4) to the guest user system A (post-5). Decrypted session key KS value of R ensures the authenticity of the info-communications systems.

Algorithm of authenticate process (fig.3)**:**

1) B: RB →:A ;
2) A: R, RS → :T ;
3) T: C(B)=E_{KTB}(RS,R,KS), &
& C(A)=E_{KTA}(R,KS,C(B)); T: C(A)→ :A;
4) A: C(B)=De_{KTA}(R, KS, C(B)) → :B
5) B: (RS, R, KS)=De_{KTB}(CB); C(A)=KS(R) → :A .

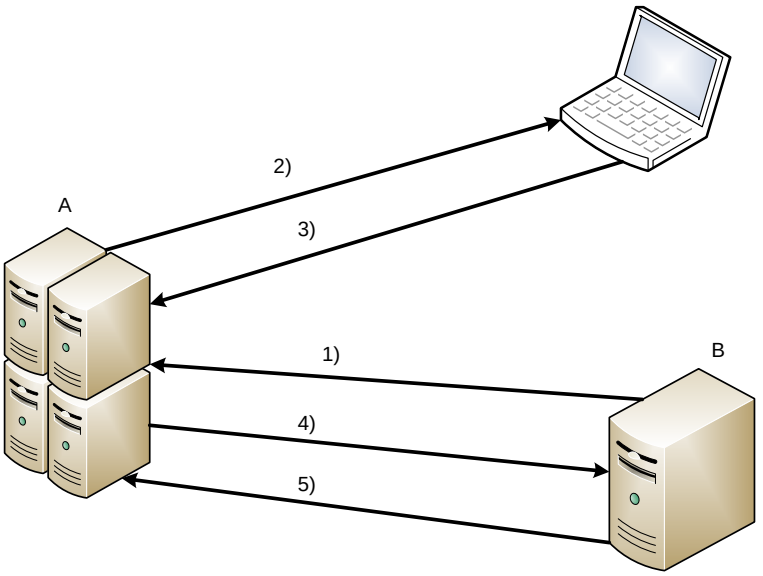


Fig.3. The scheme of mutual authentication

The overall authentication scheme as follows (fig.4).

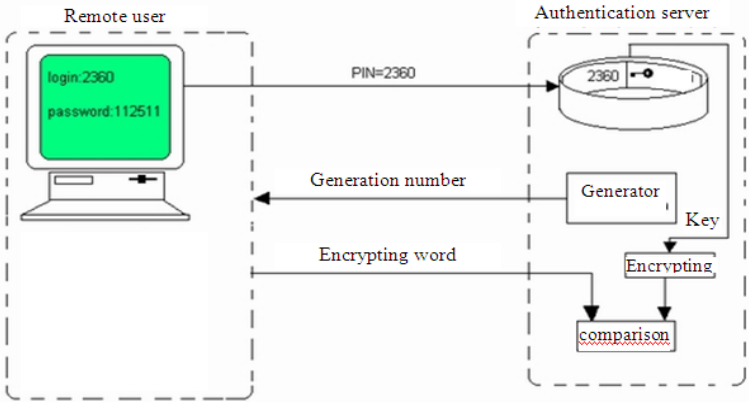


Fig.4. Scheme of Authentication process

Mechanism of generation number. The generation mechanism has its limitations - it usually requires a computer at each end of the connection, since the dongle must be allowed to receive and send information. A timing diagram allows us to restrict a simple terminal or fax. In this case, users can even enter your password to the phone keypad, when they call in the network for voice mail.

Scheme of "challenge-response" scheme gives timings for ease of use. For logic input timing diagram for the user to simply dial 10 digits. Scheme as a "request-response" may require the user to perform a larger number of manual steps. In some schemes "challenge-response" user must enter a secret key and then typing on the keyboard obtained using a hardware key, the encrypted word call. In some cases, the user must perform a second logic input to the communication server after authentication.

In the system of comprehensive protection of information resources Internet portal essential point is the application of institutional, organizational and technical measures. In particular, when you access to the Internet portal in action (connecting to the Internet) need to develop all the organizational and normative-technical documentation, including a part of information security. You need to assign and prepare the security administrator, to train other staff (content managers and editors) the basic rules of safety information. Also, to improve access to information resources Internet portal should make full use of technologies and redundancy and duplication (for example, through the use of cluster solutions) designed to maximize backup and recovery, and other measures.

5. Comparison of authentication methods

The generator of number creates one-time password (OTP) for user to have the following advantages and disadvantages.

Advantages:

1. The applications that users employ do not have to be changed, easing the implementation of OTP.

2. These passwords are typically secure from password-cracking programs because of the nature in creating these random passwords. However, if the hacker can guess

the secret key that was used to generate the listof passwords, there is a chance that he can determine the OTPs that were generated.

3. OTP defeats eavesdropping attacks. Even if a hacker sees the password, it is too late to use it because the user is authenticated and the password becomes invalid.

4. If a hacker is lucky enough to guess a randomly generated OTP, he is granted access to the account one time; subsequent access requires the hacker to get lucky again guessing a randomly generated OTP.

Summary comparison analysis of authentication methods is shown next table:

Methods of authentication	**Advantages**	**Disadvantages**
User ID and Password	1. Least expensive authentication method to use. 2. No need to carry any extra hardware device. 3. No need to install extra software. 4. User IDs and passwords can be changed at the user's choice, and most users know how to use them.	1. Weak and susceptible to numerous attacks. 2. Security depends on the users' ability to maintain the user ID and password secret. 3. Not fully reliable when used for making financial transactions remotely, such as fund transfers and bill payments through an Internet banking channel. 4. Cost of support increases with user ID and password complexity (i.e., help support or IT staff may need to spend extra time dealing with authentication problems, such as helping staff reset passwords that are locked after a certain number of failed entry attempts).
Hardware Tokens	1. More secure to use than user ID or passwords. 2. Enhance the image of the organization by securing user credentials more effectively. 3. Users don't need to remember complex passwords. 4. Can be used for login and transaction authentication.	1. Involves additional costs, such as the cost of the token and any replacement fees. 2. Users always need to carry the token with them. 3. Users need multiple tokens for multiple Web sites and devices. 4. Does not protect fully from man-in-the-middle attacks (i.e., attacks where an intruder intercepts a user’s session and steals the user’s credentials by acting as a proxy between the user and the authentication device without the user’s knowledge).
Authentication method based on the random number generator	1. OTP values are valid for a short period of time. 2. The OTP value can be simply read off the screen. 3. An attacker would need undetected physical access to the device 4. The OTP value is generated at the user’s request; no value change after a short amount of time	1. OTP values can be obtained easily by a bystander. 2. The OTP value may change while it is being entered. 3. An OTP value is valid until a new OTP value is used. 4. The user must press a button to generate the OTP value

5. Conclusion

Suggested for implementation of complex organizational and technical solutions to ensure the protection of information resources developed and already deployed public and corporate Internet portals. By providing the necessary level of information security by preventing and blocking specific threats is achieved stable operation of hardware and software PIR and the Internet portal as a whole.

Reference

1. Richard E. Smith. Authentication: from passwords to public keys. http://all-ebooks.com/
2. Smith, Richard, “Deciphering the Advanced Encryption Standard,” Network Magazine 16, no. 3(March 2001), pp.96-101.
3. J. Allen, The CERT® Guide to System and Network Security Practices, Addison-Wesley, Boston, MA (2001).
4. M. Bellare, R. Canetti, and H. Krawczyk, "Keyed Hash Functions and Message Authentication," Advances in Cryptology—Proceedings of CRYPTO '96, pp. 1–15 (1996).
5. M. Abadi, "Explicit Communication Revisited: Two New Attacks on Authentication Protocols," IEEE Transactions on Software Engineering 23 (3), pp. 185–186 (Mar. 1997).
6. Matt Bishop. Computer Security: Art and Science. Addison Wesley. November 29, 2002, ISBN:0-201-44099-7, p.1136.

About of One Method of Synthesis of the Structural Protected Computer Network

Salim Karimovich Ganiev, Madjit Malikovich Karimov, Durdona.Yakubdjanovna Irgasheva

Tashkent University of Information Technology, Tashkent, Uzbekistan,
dr.mmkarimov@rambler.ru, durdona.ya@gmail.com

Abstract. This paper used the method of synthesis of the structure of a corporate computer network, based on the evaluation of this method are given in the design of security metrics protected corporate network at different levels of the model Open System Interconnection.

Key words: synthesis, computer network, metrics of security, OSI model, structural protection.

1. Introduction

For definition of services and mechanisms of safety and application in appropriate levels the method of synthesis of structure of computer networks on the basis of the reference OSI model is offered to use them [1,2]. When ensuring information security of the corporate computer network (CN) sets of services of safety in which various functions of ensuring information security, such as confidentiality, integrity and availability are realized are developed. Services differ from safety mechanisms which are concrete measures of realization of these services. Only the protected structure is capable to make effective association of services, to provide controllability of corporate CN, its ability to develop and resist to new threats at preservation of the set properties. Therefore at creation of difficult corporate CN it is necessary to consider a number of the specific conceptual requirements directed on safety of functioning of corporate CN. These requirements are created on the basis of the concept of safety [1,2] and include five main components:

1. The definition of security services.
2. Definition of security mechanisms.
3. The layered model for building security services.
4. Correlation of the security services and the level model.
5. Correlation of security mechanisms and services.

2. Services of security in OSI model

The most important element of this standard is the definition, what services of safety have to be provided at each level of the reference OSI model. Distribution of services of safety is given in work [1] on levels of reference model. Using various services of safety and the corresponding mechanisms of their realization in various levels of the OSI model, it is possible to project corporate CN with various levels of safety. Depending on number and the structure of services of safety realized in corporate CN, it will have various degree of security. Naturally, in a design stage of corporate CN influence of separate services on information security of corporate CN isn't known. Therefore to express this dependence in a general view, we will use the equation of multiple linear regression:

$$M_S = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k \quad (1)$$

where $\alpha_0, \alpha_1, \ldots, \alpha_k$ - coefficients of multiple linear regression; $x_0, x_1, \ldots, x_k$ - the parameters characterizing services of safety; M_s - metrics of an assessment of information security of CN (where, S - set of services of safety ($S = \{s_j\}, j = \overline{1,k}$) and $0 < M_s \leq 1$ - the higher the value M_s , the higher the security level corporate CN).

The value $x_0, x_1, \ldots, x_k$ can be determined using a model of reliability offered by H.Pham. Multiple linear regression coefficients $\alpha_0, \alpha_1, \ldots, \alpha_k$ also determined on the basis of statistical data on the violations of information security in an enterprise of the CN, where - α_1= 0,03, α_2= 2,65, α_3= 13,72, α_4=15,11. Denote $\sigma_i, i = \overline{1,\rho}$, the correction factors characterizing security services at different levels. In this $\sigma_1 < \sigma_2 < \ldots < \sigma_\rho$ and $0 < \sigma_i \leq 1$. Here ρ - the number of levels (according to the model OSI ρ=7). Then, taking into account the correction factor equation (1) can be written as follows:

$$M_s = \alpha_0 + \sum_{j=1}^{k} \sum_{i=1}^{\rho} \alpha_j x_j \gamma_{ji} \quad (2)$$

where:

γ_{ji} = { σ_i, *if j – y service of safety is used on i – m level*; 0-*otherwise;*

Thus the condition has to be met $\sum_{i=1}^{\rho} \gamma_{ji} = 1$

This condition means that at each level only one correction coefficient is used. Thus, the formula (2) allows to calculate numerical value of level of information security and level at which these means are applied. Besides, using the same formulas, it is possible to solve the return problem, i.e. to project CN with the set level of safety. In this case the task is formulated as follows:

$$\alpha_0 + \sum_{j=1}^{k} \sum_{i=1}^{\rho} \alpha_j x_j \gamma_{ji} \geq M_0$$

where M_0 – predetermined level of security.

Currently CN building, usually on the basis of the model TCP/IP. This model has four levels (Table 1) and the correction factor of four. In this $\sigma_1 < \sigma_2 < \sigma_3 < \sigma_4$ and $0 < \sigma_i \leq 1$, $i = \overline{1,4}$.

3. Example of synthesis process of protection

Safety at transport level allows to supervise process of establishment of virtual connections and information transfer on them, and functioning at applied level is capable to provide the most reliable protection. Having analysed OSI model levels, it is possible to draw the following conclusion that the management of access functioning at applied level, is capable to provide the most reliable protection. Considering this fact, as an example we use

the following correction coefficients at synthesis of system of protection for corporate CN: σ_1=0,25, σ_2=0,5, σ_3=0,75, σ_4=1 (table 2).

Table 1.

Correction coefficients in the TCP/IP model

The levels of the OSI model	TCP/IP model	Correction coefficient
7. Applied level	Applied level (HTTP, SMTP)	σ_4
6. Representative level		
5. Session level		
4. Transport level	TCP	σ_3
3. Network level	IP	σ_2
2. Channel level	Channel level	σ_1
1. Physical level		

Table 2.

Information security mechanisms

Levels TCP/IP model	Mechanism	Correction coefficient
Applied level	Access control	1
Network level	Access control	0,75
Transport level	Access control	0,5
Channel level	Routing control	0,25

To assess the degree of information security of the projected corporate KS use the following procedure. On the basis of the known statistics on violations of information security (α_1= 0,03, α_2= 2,65, α_3= 13,72, α_4=15,11), the following expression for the metric of security:

$$M_s=0,03x_1+2,65x_2,+13,72x_3+15,11x_4 \quad (3)$$

where x_1, x_2, x_3, x_4 – parameters characterizing the confidentiality, integrity, availability and authorization, respectively. Given the correction factor equation (3) is:

$$M_s=0,03x_1\gamma_1+2,65x_2\gamma_2 ,+13,72x_3\gamma_3+15,11x_4\gamma_4 , \quad (4)$$

Since routing management provides privacy and authentication at the link layer is in this case $x_2=0$, $x_3=0$, γ_2 =0, $\gamma_3=0$. Then equation (4) becomes:

$$M_s^{(chan)}= 0,03x_1\gamma_1+15,11x_4\gamma_4 \quad (5)$$

Thus, using the routing control data link level security metric expression is (at $\gamma_1=\gamma_4=0,25$):

$$M_s^{(chan)}=0,03x_1\bullet 0,25+15,11x_4\bullet 0.25=0,0075x_1+3,77x_4 \quad (6)$$

Using data: $x_1=0,697$ и $x_4=0,031$, from (6) we will receive final result of $M_s^{(chan)}=0,13$.

When using management by access at network level, it agrees (5), expression for a metrics of information security will look like (at $\gamma_1=\gamma_4=0,5$): $M_s^{(net)}=0,0134x_1+7,56x_4$, and taking into account numerical $x_1=0697$ and $x_4=0,031$ we will receive $M_s^{(net)}=0,2436$.

At management of access at transport level, it agrees (5), expression for a metrics of information security will look like (at $\gamma_1=\gamma_4=0,75$):

$$M_s^{(tr)}=0,02x_1+11,3x_4 \quad (7)$$

and taking into account numerical values for a metrics of information security (at $\gamma_1=\gamma_4=0,75$) it is received: $M_s^{(tr)}=0,37$.

For applied level, using service of management by access (at $\gamma_1=\gamma_4=1,0$) taking into account numerical x_1=0,697 and x_4=0,031 is received the following result:

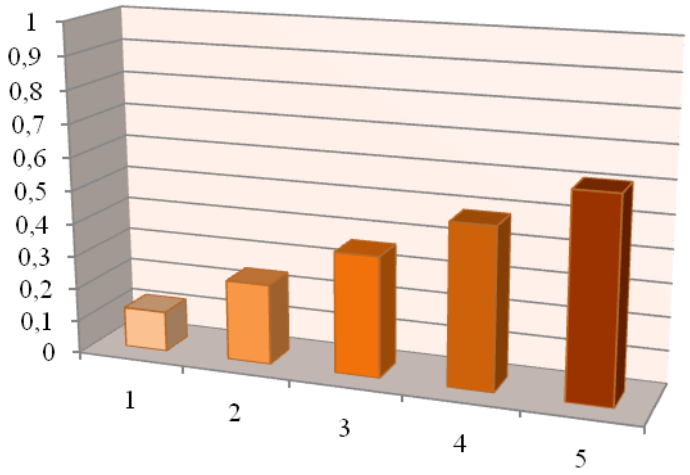
$$M_s^{(app)}= 0,027x_1+15,11x_4 =0,49 \quad (8)$$

Now, we will consider design of system of safety for corporate CN wth use of service of management by routing at channel level and managements of access at applied level. From above given calculations follows that the metrics of an assessment of safety at channel level is equal $M_s^{(chan)}= 0,13$, and at the applied $M_s^{(app)}$=0,49. If to unite these two services it is possible to receive result of the general assessment of a metrics of safety in the following look: $M_s=0,13+049=0,62$. The received results show that sharing of the mechanism of management by access at channel and applied levels (tab. 3) is most expedient.

Table 3.

Results of security metrics

TCP/IP model levels	The routing control	Access control	The combined service
Channel level	0,13	-	-
Network level	-	0,2436	-
Transport level	-	0,37	-
Applied level	-	0,49	0,62



Evaluation metrics security of corporate networks

4. Conclusion

The figure shows the evaluation metrics safety in the design protected corporate network based on the routing control mechanisms and access control at different levels of the model TCP / IP. In this one-column means that by using the routing control mechanism of the link layer can be marginally effective from a security standpoint. 2 -, 3 -, and 4 - columns show that the use of access control mechanism at the network, transport and application levels, respectively, gives a higher score. But it is advisable to use the combined mechanisms (column 5), which are at different levels of the model TCP/IP.

The results of synthesis of structural - protected corporate network have shown the feasibility of joint use of access control mechanisms at the data link and application levels.

References

1. Alguliev R.M., Ragimov E.R. A method of estimating the information security of corporate networks in the design stage//Information Technology. 2005 - № 7.
2. ISO/IEC 17799:2000 Information technology - Information security management.
3. ShirochinV.P., MukhinV.E.,KulikA.V. Design Issues of information security in computer systems and networks. Kiyev; «VEK+». 2000. – 111 p.
4. KovalevskiyV. Cryptographicmethods.// Computer Press, № 5, 1993.
5. GanievS.K., KarimovM.M. «An integrated system of information security in corporate networks»./Theses of reports of the International SCP «Innovation 2001».-Tashkent, 2001.-P.97-98.
6. V.A. Serdyuk. Analyze of modern lines of construction of models of information attacks// Information technologies, № 4, 2004.

The New Development Project of an Integrated Information System of Management in the Todor Kableshkov University of Transport

Dimitar Dimitrov

Todor Kableshkov University of Transport, 1574 Sofia, Bulgaria
ddimitrov@vtu.bg

Abstract. This paper presents a new project for the development of an integrated management information system at the Todor Kableshkov University of Transport. The main objective of the project is to achieve the complete integration of all processes by building an integrated information system encompassing all functions and systems. The implementation of the project is based on established contemporary trends in web development.

Keywords. web integrated information system, information services, management.

1. Introduction

The construction of contemporary information systems is a complicated task, which is related to the management of an increasing number of human activity processes. One of the main objectives of these systems is to improve information servicing of all consumers, which comprises the following compulsory characteristics:

-modularity and coordination of various modules;
-use of a common data base;
-automated binding with other systems and data bases;
-real time work;
-high degree of data security in the process of systems' operation.

The design and use of such systems is a complex task, which involves a series of hardships as to the adequate definition and description of physical processes and the solution of a number of specific problems.

The paper presents the results of a current project on the construction of a system of unified management of processes in the Todor Kableshkov University of Transport. The basic objective is the complete integration of all processes in the Todor Kableshkov University of Transport by means of a unified information system construction, aimed at embracing all functions and systems within the so-called Integrated Information System of Management (IISM).

2. Main University Processes and Target Groups

Main University activities ate connected to education process management as well as to the planning and carrying out of tasks of corresponding categories of personnel, programmes and processes [1]. Figures displayed below show typical patterns of movement and interaction between separate categories (groups) of personnel and resources, target of automation and information assurance.

Fig. 1 shows the University itinerary of three basic target groups of people (students, lecturers, administrative and service personnel). Fig. 2 illustrates the triangular structure of basic target groups' interaction, and fig. 3 represents the process of planning and management of University activities and resources.

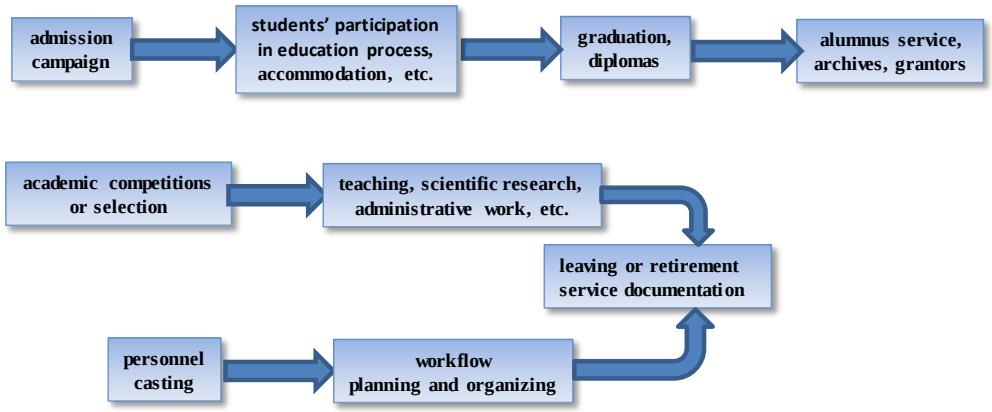


Figure 1. Basic groups' University itinerary

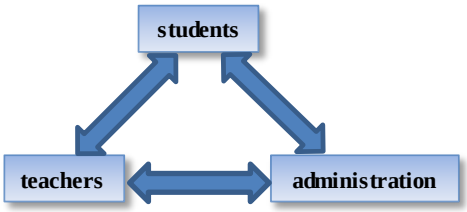


Figure 2. Interaction between basic University groups

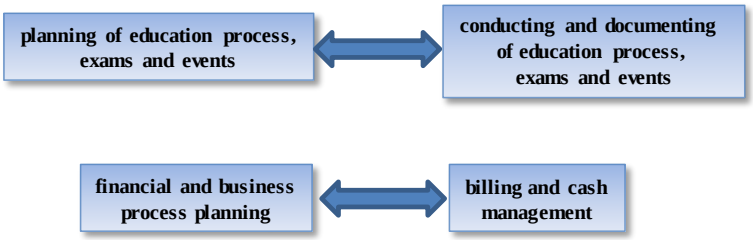


Figure 3. General representation of the process of planning and management of University activities and resources

In compliance with various target groups' needs as well as with their University itinerary, basic screens and data base structure as to these groups' information assurance have been developed in detail [2]. Methodological aspects and algorithms of change of states and registration of events, optimization of planning and of education and work tasks reporting system, have been also developed. The general mathematical model of the integrated information system is represented below.

3. A Mathematical Model of IISM

The mathematical model expresses separate sets and categories of participants in all University processes. The model represents events and activities of various categories of University participants. The main active category is represented by the set of students, subdivided into three types, and namely: students, post-graduate students and PhD students. Specific services are provided to all members of students' set, these services being related to education process, conditions of life and administrative attendance, as shown in formula 1.

$$s_{i,t} \in S \rightarrow f\{ cur_{i,t}, exmn_{i,t}, tax_{i,t}, acmd_{i,t}, scsh_{i,t}, perst_{i,t} \}, \quad (1)$$

where
S – all students at University;
$cur_{i,t}$ – curriculum;
$exmn_{i,t}$ – examinations;
$tax_{i,t}$ – taxes;
$acmd_{i,t}$ – accommodations;
$scsh_{i,t}$ – scholarships and bonuses;
$perst_{i,t}$ – period of study;
$s_{i,t}$ – student i of type t. The type t is *(1 - Student,2 - Postgraduate Student, 3 - PhD Student).*

The other essential group is represented by teaching staff performing activities related to their educational and administrative commitments and services, which is expressed by means of formula 2.

$$t_j \in T \rightarrow f\{ tctp_j, aatg_j, othact_j, pmnt_j \}, \quad (2)$$

where
$tctp_j$ - training courses and teaching process;
$aatg_j$ - appraisal and academic titles promotion;
$othact_j$ – other activity
$pmnt_j$ – payments;
t_j – teacher j.

Other University personnel activities and services are taken into account in formula 3.

$$p_k \in P \rightarrow f\{ act_k, pmnt_k \} \quad (3)$$

where
act_k – personnel activity;
$pmnt_k$ – payments;
p_k – personnel k.

Optimality achievement in the capacity of separate groups of sets leads to the optimality of overall University work capacity, in accordance with formula 4.

$$f\left(S_{opt}, T_{opt}, P_{opt}\right) \rightarrow OWPU \quad (4)$$

where

OWPU – Optimal Work Performance in University.

Represented general model reveals basic services and activities carried out in a university organization in the purpose of satisfying the needs of separate categories of groups and sets. The model can be complemented on a modular principle in accordance with concrete specific characteristics.

4. System Deployment Environment

The designed information system is expected to entirely function online and to be founded on contemporary web visualization and work tools. Server platform is based essentially on open source operation medium Linux and Java programming platform, whereas data base uses a free Oracle platform version, as, according to current project, resources expected to be used fit into the framework of this free data base environment.

Integration with the other operating systems is expected to be carried out by means of synchronizing used data bases through convenient automated interfaces and communication tools. The implementation of the so called *electronic banking*, which is going to bring about complete planning and management of cash information flows and transactions, is also anticipated.

As far as data security is concerned, multi-consumer work with the system is envisaged with individual rights and roles, including in a medium with base security certificates of an incorporated electronic certificate for the whole University domain.

5. Conclusion

Designed integrated information system is constructed on a modular principle comprising many advantages. In the first place, this is the single input of information and its numerous multi-purpose uses, as well as its accessibility anywhere and anytime, when needed. This feature facilitates and improves the activity of Education Department and Deans' Offices through processes automation and issuing of required documents. This will certainly lead to maximum visualization of all ongoing University processes as well as to more efficient use of time, resources and premises.

The Information system envisages the complete integration of Learning Management Systems (LMS), virtual library, documents turnover system, tangible resources management and University accounting system [3]. All this is a prerequisite for a more efficient and easier tracing of financial flows and transactions, as well as for the management of incomes and costs, including strategic planning assistance.

In a technical sense, the system is aimed at eliminating recurrent actions in communication process, at information loading and documents turnover managing as well as at minimizing the use of paper documents, files, and etc. Another important objective is the fast and exact preparation of information queries, reports, analyses and surveys as to students, lecturers and their professional activity. The system's model is based on a web-oriented platform, which excludes the need of consumers' personal attendance in the case of most services, including by means of electronic banking and management of financial resources.

References

1. Dimitrov D., Vasilev D., Functional Analysis Of Newly Integrated Information System Of Management of Todor Kableshkov University Of Transport, Conference proceeding of 21st International Scientific Conference ТРАНСПОРТ-2013
2. Technical Documentation of Project BG051PO001-3.1.08-0007, „Enhancement of Education Process Efficiency through Perfecting the Management Systems of Quality of Services in Education and of Training at the Todor Kableshkov University of Transport", Operative Programme "Human Resources Development", co-financed by EU European Social Fund.
3. Technical Documentation of Project BG051PO001-4.3.04 „Development of Distant Education Electronic Forms in Higher Education System", Operative Programme "Human Resources Development", co-financed by EU European Social Fund.

Simplified On-line Approach for Proper Retort Heating Profile Determination in the Food Canning Industry

Nanko Bozukov[1], Hristo Dinkov[2], Marko Dimitrov[3], Vasil Shikov[4]

[1] UFT, 4002 Plovdiv, Bulgaria
bozukovnanko@abv.bg

[2] UFT, 4002 Plovdiv, Bulgaria
hristo.d@dir.bg

[3] Energy rating company- Zenit Ltd, 1700 Sofia, Bulgaria
keozenit@gmail.com

[4] UFT, 4002 Plovdiv, Bulgaria
vshikov@yahoo.com

Abstract. This paper presents an easy practical approach to the solution of common commercial situations in the food and canning industry that need prompt decisions. The use of information equipment and technologies that are strongly prescribed for application by international laws and standards such as ISO and HACCP facilitates changes in the process conditions and time-temperature profile aimed at maintaining the technological values involved in food safety and energy efficiency. This study provides an example from the food sterilisation industry. At times, due to customer or market demand, manufacturers need to vary the type, material and volume of product packages. The different thermal penetration properties lead to the need for a prompt reaction towards suitable changes in the scheduled time-temperature process in order to preserve the important, scientifically defined and desired accumulated lethality F_0 under the conditions of pre-defined sterilisation temperature. The necessary technical equipment, specific tools and practical algorithm have been presented. The data on the example given have been taken from a real industrial process, and the discussion of the results has been supported by table and graphic representations. Finally, conclusions and recommendations have been given concerning the conditions which should be observed in order to guarantee the certain reliability of the approach.

Keywords: Data, Excel, temperature measurement, 1-Wire buss,

1. Introduction

According to the dynamic changes in market demands, the producers of canned food are often forced to change their products – their volume, shape and the material of the containers. This always differs from the familiar and appropriate manufacturing practices concerning the time-temperature schedule and creates an urgent need for correction. The lack of practical experience in using new packages with different thermal penetration characteristics and changed volume makes that process relatively complicated. An adequate heating process is essential: if insufficient, it can affect the safety or shelf life of the product; if overextended, the product quality will deteriorate and energy losses will be incurred. On the other hand, food safety demands and good manufacturing practices (ISO and HACCP systems) force producers to implement and use computer-based measurement and backup systems in their industrial practice [1]. In any case, the availability of an appropriate validated on-line measurement system makes the above task much easier [1, 3]. Having the necessary data, following a simple algorithm of steps, and using Excel for data acquisition, it is fully possible to determine a new time-temperature heating curve which will both cover

the demands of the specified food safety (target lethality), and reduce undesirable over-processing and energy losses.

1. Equipment, materials and methodology.

In our study, we accepted the following initial conditions:

- On-line measurement and file recording system in use;
- Experience in thermal sterilisation processes in the manufacture of certain products and packages;
- Qualified food technologist and retort operator;
- Specially designed device for on-line measurement of the so-called "cold spot" within the package placed in the retort package container, and a measuring device for the retort temperature.

Let us suppose that because of customer demand, the package needs to be changed to a relatively bigger or smaller one. Evidently, this situation will make it necessary to re-adjust the processing conditions as a result of the different thermal penetration and geometrical characteristics of the package, and the quantity of the raw material filled.

2.1. Equipment and materials needed:

A. PC-based measurement and recording 1-Wire bus system used under industrial conditions.

Apart from the standard Windows-based software, software for on-line visualisation and calculation of accumulated lethality (F_0 value at the cold spot) was defined, designed and used for that purpose. Briefly, we can define the function ability and main characteristics of this software module as follows:

- Dynamic table generation for all sensors connected to the system;
- A daily, weekly or another time interval report, either in table or graphic format;
- Document printing ability;
- Database storage (in text format);
- On-line view of the selected temperature points, and calculation and graphic presentation of the accumulated lethality (F_0-value). (when possible)

In this case, we will not discuss the structure and hardware ingredients of the system because any system which has the above abilities can serve this purpose well.

Having in mind that the data is recorded in text format, it is easy for any experienced employer to import and process the collected data in Excel, and then present them graphically.

B. Industrial Retort

C. Familiar, already manufactured canned food, the processing parameters of which have to be achieved under the new conditions: changed shape, material and/or volume of the package (container). Standard products such as peas and tomato paste were used in the examples given in the article.

2.2. Algorithm and step sequence of the approach for defining new and actual heating and cooling temperature retort profile:

1. Placing the measuring device into the jar or another package at the place corresponding to the so-called "cold spot", and of the package so prepared – at the coldest spot position of the retort rack (basket).
2. Defining the target F_0 value (accumulated lethality) – known for almost all of the processed products.
3. Defining the proper temperature profile for the retort (the usual one for the product).
4. Starting the sterilisation process and controlling the process evaluation – retort and product temperatures and F_0 value on the monitor. *(when it is possible)
5. Starting the cooling procedure after reaching about 90% of the target F_0 value. *(when it is possible)
6. Importing the process data into an Excel prototype sheet where all results are recalculated and graphically presented.
7. Making a decision about the process duration comparing the obtained and the target F_0 value.
8. Simple correction of the time of the sterilisation temperature (the relatively constant part of the profile), extracting or adding minutes (corresponding to the exact F value within that interval of minutes we wish to change), until the target value is reached.
9. Checking the result under the newly defined conditions. Steps 6, 7 and 8 may be repeated in order to obtain the final time-temperature schedule that has to be followed during the sterilisation process.

* In cases where the control system gives the opportunity for on-line F_0 calculation.

3. Discussion and results

In both cases discussed below we suppose that the measurement system does not have the ability to calculate F_0 on-line. That mince the above proposed steps 4 and 5 are fulfilled as following time-temperature profile defined in step 3.

In accordance with the proposed algorithm, the data on the sterilisation process of tomato paste in TO 0.250 jars processed under the conditions known as suitable for bigger packages were imported into Excel. The calculated data are graphically presented in Fig. 1. The data showed that the product was over-processed: F_0 was equal to about 60 minutes.

The F_0 value was calculated for corresponding temperature and time at 1-minute intervals. As reported in many scientific sources, the sufficient F_0 target value is about 25 minutes, which may differ slightly depending on the country demands, according to a base temperature of 93.3 ^{0}C and z=8.8 [2]. By reducing the heating time and extracting the corresponding F_0 values, we can reach the desired value. For correction purposes, the rest of the values for package temperature were simply shifted to the place of the missing positions in the table, relying on the same velocity of heating and a relatively constant F_0 value in the cooling phase. The result is shown in Fig. 2.

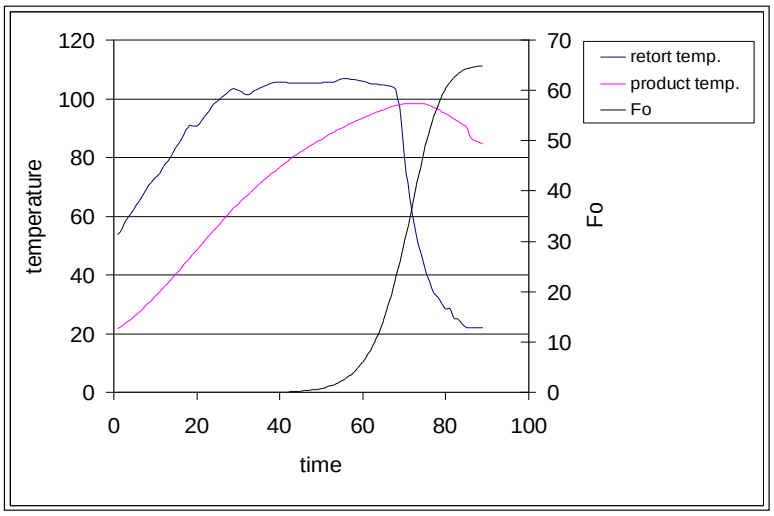


Fig.1 Tomato paste sterilisation process

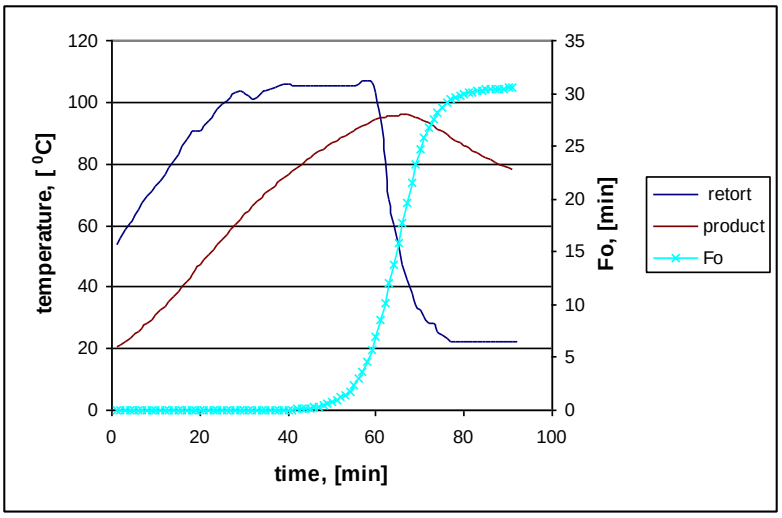


Fig.2 Sterilisation process after correction

The example shown above demonstrates the proposed algorithm in the case of over-processing followed by shortening of the cooking time in order to reduce the energy losses and quality deterioration. This approach was tested again for different products under industrial conditions. The next example shown in Fig.3 illustrates the process of green pea sterilisation, where the heating process was insufficient with regard to food safety (the F_0 value was 3.8 instead of the recommended approximately 6.2 conditional minutes).

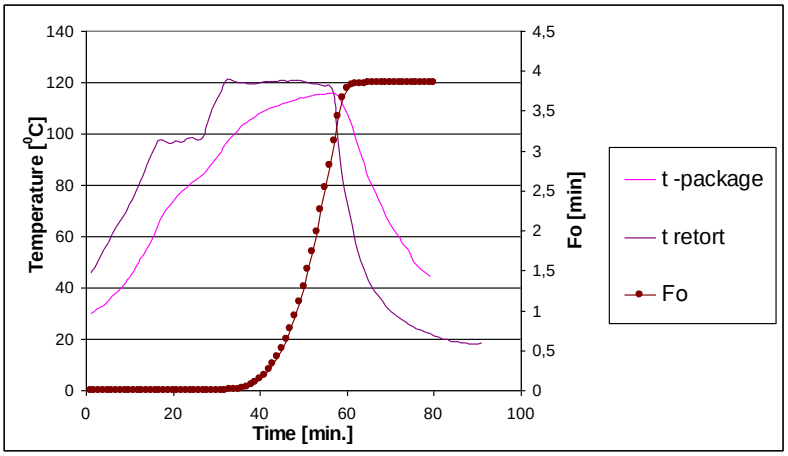


Fig.3 Pea sterilisation without correction

In this case, we extended the heating time interval where the temperature was relatively constant and added the calculated F values for the maximum temperature reached at the cold spot of the product until the target F_0 value of about 6 minutes was reached. We need to be fully aware that the time-temperature curve we define will practically yield a greater F_0 value under the same conditions. This is due to the type of heating process and the fact that the product temperature continues to increase during that time. The inconvenience and inaccuracy of the method can be overcome by a repetition of the same method, followed by another correction, instead of process modelling, which is possible but complicated. If necessary, having the real data, we can adjust the schedule by extracting values and repeating the steps (similarly to the first example, i.e. Fig.1 and Fig.2). The final result is presented in Fig.4 and illustrates the practical effect of the corrected process when the steps of the proposed procedure are followed.

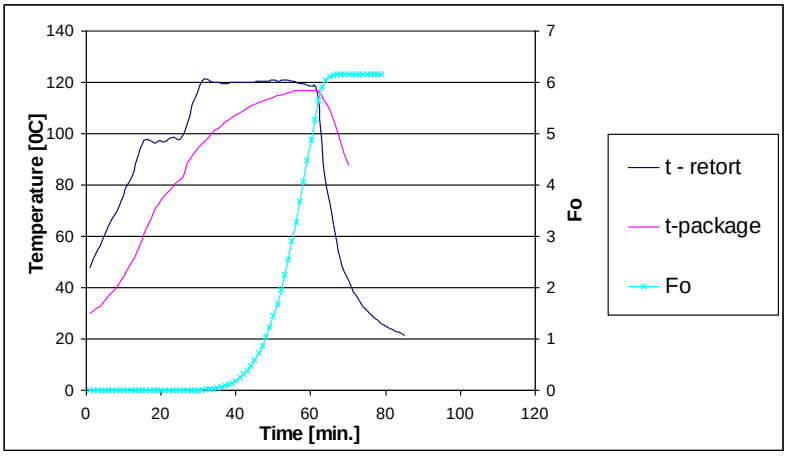


Fig.4 Practical result of the time-temperature curve correction

4. Conclusion

The algorithm is obviously applicable even in cases where the measurement system is not able to calculate and present on-line the values of temperatures and F_0.

Experiments with other products replicated the expected results and the observed accuracy remained within 10% of the target F_0 value. In our opinion, for express practical solutions, the proposed approach is simple and easily applicable for process correction in order to meet the demands for safe production and energy efficiency.

Overall, any process deviations concerning the initial product temperature or the time-temperature profile of the retort lead to mistakes. The proposed approach is reliable when the processing conditions are stable and repeatable. It may be used for process condition and temperature profile definition but will not guarantee the result in the event of process deviations.

References

1. Simpson R., A. Teixeira, S. Almonacid, Advances with intelligent on-line retort control and automation in thermal processing of canned foods. *J. Food Eng*.,vol. 18, July 2007, p.821-833
2. Хайдутов М., Н. Пенов, Ст. Делкинова. Режими на стерилизация , F_0 – стойности и тяхното влияние върху безопасността, качеството и енергоемкостта на стерилизирани консерви., Пловдив,Академ. изд. на УХТ, 2005 г.
3. Simpson R., I. Figueroa, D. Llanos, A. Teixeira, (2007). Preliminary validation of on-line correction of process deviations without extending process time in batch retorting: Any low-acid canned foods, Food control, vol. 18, Issue 8, p.983-987
4. Христо П. Динков, Системите за температурен контрол и архивиране на данни по отношение управлението на качеството и безопасността на храните., Научни трудове на УХТ- Пловдив, Том LIX-2012, стр.694-696
5. *Х. П. Динков, П. Т. Панайотов, П. И. Никовски, Р. Х. Динкова,* Приложение на информационно-архивиращите системи за оптимизиране на режимите и снижаване на енергийните разходи при топлинна стерилизация на храни., Научни трудове на УХТ-Пловдив, Том LIX-2012, (стр.694-696)
6. Norohna, J., Hendrickx, M., Van Loey, A., & Tobback, P. (1995). New semi-empirical approach to handle time-variable boundary conditions during sterilization of nonconductive heating foods. *J. Food Eng.*, *24*(2), 249–268.

Numerical Procedure to Determine the Optimal Composition of the Steel

Nikolay Tontchev, Emil Inanov, Georgi Georgiev, Ilian Lubomirov
'Todor Kableshkov' Transport University, Sofia, Bulgaria

Abstract: The paper presents a numerical procedure in order to determine the optimal composition of the steel using a small-size database containing information about composites and their properties in one and the same process. The procedure uses a multi criteria decision for specific properties of the investigated materials, necessary for the design formulated via a generalized parameter of the chemical composition of the alloy. The study is carried out with various linear combinations of generalized parameters of the chemical composition and consequently determining the optimal quantities of the individual components of the composition.

Keywords: Steels, determine the optimal composition, optimal mechanical properties

1. Introduction

In the optimization of the composition and properties of traditionally used or new alloy compositions in most cases there is limited preliminary information about the research grade material under appropriate thermal and mechanical treatment. Therefore it is necessary to create methods and approaches, giving a rational result for a limited number of data, determined by the composition and characteristic important design properties.

Approaches developed by us (Tontchev (2007,2013), Koprinkova (2011,2013)), characterized by the use of large amounts of information via modeling of neural networks and regression analysis, can be used to verify the decision in the case when the explored class is added to a large base of materials.

The purpose of this exposure is to propose a numerical procedure in order to determine the optimal composition of the alloy using a small volume database for composites and their properties in the same process. The procedure is based on determining a multicriteria decision for characteristic of the studied material properties, expressed by generalized parameters of chemical composition.

2. Main parts

2.1. Nature of the Procedure

The advantage of the procedure is the use of a small-size database which can be determined on the basis of preliminary tests. Reducing the number of observations in the sample is due to the reduction of the problem to a generalizing parameter of the chemical composition. Such a sample size of the observations in conventional tests with more than five elements is not sufficient for the tools of the method of approximation with a neural network and with regression analysis. This is directly related to the accuracy of the solution.

The solution of the problem is reduced to determining the single-objective models, fixing the influence of the studied properties of the generalized parameters of the chemical

composition of the alloy by standard approximations. The choice of fixed approximation is determined based on the highest coefficient of multiple correlation R. Formulated models are verified using F - Fisher criterion. Algorithm of the procedure is represented in Fig.1.

Problems of materials science, as problems for optimization of the composition and properties, are problems for multicriteria optimization. A multicriteria problem for materials science include at least two criteria – criteria of contradictory nature, e.g. strength and ductility. In such problems, it is possible to determine the optimum alloy composition, corresponding to the preset requirements. On the basis of formulated models, after the approximation the multicriteria problem is defined with all requirements defined in it. Identical values of the tested criteria should apply to all linearly independent equations in the system of generalizing parameters of the chemical composition. The peculiarity of the problem is that it is reduced to determining the values of the generalized parameters of the chemical composition under these fixed conditions. This condition ensures the solution of the problem of optimal composition after fulfilling the requirements' criteria.

2.2. Defining the Procedure.

The procedure was chosen to test a specific Ukrainian database for steel in Energomash listed in Table 1. At the defined task, the goal is an optimal composition of steel with yield strength Re > 500 MPa and resistance to brittle fracture. Resistance to brittle fracture is achieved by using materials with relatively high values of elongation. In the practice for the considered class of steels, the optimal combination of the yield strength (Re) and the elongation (A) is respectively Re > 500 MPa and A > 11%. This requirement is defined as the minimum threshold of the multicriteria optimization problem.

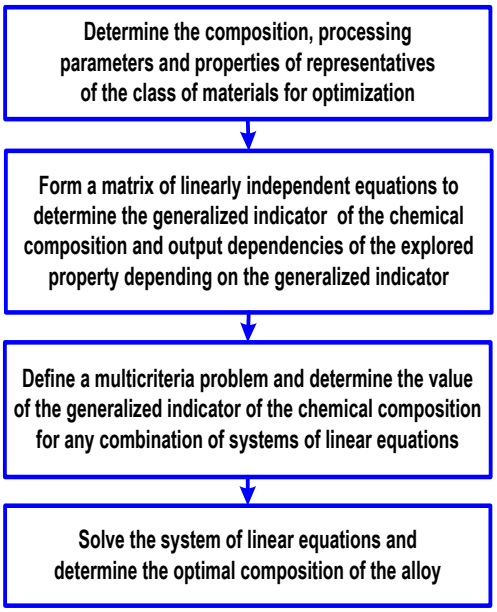


Fig.1. Algorithm of the procedure

Problems of materials science, as problems for optimization of the composition and properties, are problems for multicriteria optimization. A multicriteria problem for materials science include at least two criteria – criteria of contradictory nature, e.g. strength and ductility. In such problems, it is possible to determine the optimum alloy composition, corresponding to the preset requirements. On the basis of formulated models, after the approximation the multicriteria problem is defined with all requirements defined in it. Identical values of the tested criteria should apply to all linearly independent equations in the system of generalizing parameters of the chemical composition. The peculiarity of the problem is that it is reduced to determining the values of the generalized parameters of the chemical composition under these fixed conditions. This condition ensures the solution of the problem of optimal composition after fulfilling the requirements' criteria.

2.3. Testing in small database

The procedure was chosen to test a specific Ukrainian database for steel in Energomash listed in Table 1. At the defined task, the goal is an optimal composition of steel with yield strength Re > 500 MPa and resistance to brittle fracture. Resistance to brittle fracture is achieved by using materials with relatively high values of elongation. In the practice for the considered class of steels, the optimal combination of the yield strength (Re) and the elongation (A) is respectively Re > 500 MPa and A > 11%. This requirement is defined as the minimum threshold of the multicriteria optimization problem.

As examples of generalized parameters of the chemical composition serve used in the literature (Gorni A.A.) various dependencies for CEV (Carbon Equivalent Value or Equivalent Carbon Content). These relationships define the system of independent linear equations of the system to determine the value of the percentage of the individual ligating elements. The number of equations matches the number of components in the chemical composition of the alloy. The system has the form:

$$\begin{array}{c} \begin{matrix} C & Si & Mn & Ni & Cr & Mo & V \end{matrix} \\ \begin{pmatrix} 1 & \frac{1}{25} & \frac{1}{16} & \frac{1}{60} & \frac{1}{20} & \frac{1}{40} & \frac{1}{15} \\ 1 & \frac{1}{24} & \frac{1}{6} & \frac{1}{40} & \frac{1}{5} & \frac{1}{4} & \frac{1}{14} \\ 1 & \frac{1}{24} & \frac{1}{10} & \frac{1}{40} & \frac{1}{5} & \frac{1}{4} & \frac{1}{10} \\ 1 & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} \\ 1 & 0 & \frac{1}{6} & \frac{1}{15} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 1 & \frac{1}{30} & \frac{1}{20} & \frac{1}{60} & \frac{1}{20} & \frac{1}{15} & \frac{1}{10} \\ 1 & \frac{1}{6} & \frac{1}{6} & \frac{1}{15} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix} \end{array}$$

Table. 1: Sample database steels by Catalogue of steel grades on the composition and the properties after quenching and high temperature tempering.

№	Brand	C	Si	Mn	Ni	Cr	Mo	V	Re [MPa]	A [%]
1	40CrMnMo7	0,4	0,4	1,5	0,6	1,9	0,2	0	700	12
2	40CrMnMo58-6	0,4	0,4	1,5	0	1,9	0,2	0	830	9
3	40CrMoV5-1	0,4	1,1	0,4	0	5,2	1,35	1	1300	7
4	56NiCrMoV-7	0,55	0,25	0,75	1,65	1	0,5	0,9	1280	12
5	17NiCrMo6-4	0,17	0,3	0,75	1,35	1	0,2	0	910	15
6	18CrNiMo7-6	0,18	0,35	0,75	1,55	1,65	0,3	0	785	8
7	20MnMoNi4-5	0,2	0,3	1,25	0,6	0,5	0,5	0	420	17
8	20NiCrMo2-2	0,2	0,37	0,8	0,55	0,5	0,2	0	193	26
9	25CrMo4	0,25	0,4	0,75	0	1,05	0,25	0	650	17
10	27CrNiMoV14-5	0,27	0,14	0,3	3,7	1,55	0,35	0,1	800	11
11	26NiCrMoV11-5	0,27	0,15	0,27	2,7	1,5	0,35	0,1	730	11
12	30X2H2MA	0,29	0,37	0,45	2	1	0,4	0	700	12
13	30CrNiMo8	0,3	0,37	0,45	2	2	0,4	0	658	10
14	34CrNiMo6	0,34	0,37	0,65	1,5	1,5	0,2	0	600	12
15	34CrMo4	0,34	0,25	0,55	0,4	1,5	0,48	0	490	15
16	42CrMo4	0,4	0,41	0,25	0	0,95	0,22	0	300	14

According to the concrete equations, the corresponding generalized indicator of the chemical composition is calculated for each composition from the database / Table 1 /.

The results of these calculations are shown in Table. 2.

Table 2. Data summary indicator for the chemical composition of the sample database.

№	Cekv1	Cekv2	Cekv3	Cekv4	Cekv5	Cekv6	Cekv7	Re	A
1	0,620	1,112	1,012	0,630	1,110	0,607	1,177	700	12
2	0,610	1,097	0,997	0,600	1,070	0,597	1,137	830	9
3	0,829	1,961	1,963	0,853	1,977	0,907	2,160	1300	7
4	0,757	1,116	1,092	0,803	1,265	0,797	1,307	1280	12
5	0,306	0,591	0,541	0,350	0,625	0,303	0,675	910	15
6	0,357	0,763	0,713	0,410	0,798	0,358	0,857	785	8
7	0,338	0,661	0,578	0,358	0,648	0,341	0,698	420	17
8	0,304	0,513	0,459	0,321	0,510	0,300	0,572	193	26
9	0,372	0,664	0,614	0,373	0,635	0,370	0,702	650	17
10	0,449	0,823	0,806	0,577	0,967	0,462	0,990	800	11
11	0,428	0,783	0,768	0,524	0,885	0,442	0,910	730	11
12	0,426	0,730	0,700	0,501	0,778	0,435	0,840	700	12
13	0,486	0,940	0,910	0,561	0,988	0,495	1,050	658	10
14	0,500	0,851	0,808	0,551	0,888	0,498	0,950	600	12
15	0,478	0,872	0,835	0,499	0,854	0,490	0,896	490	15
16	0,485	0,704	0,687	0,492	0,676	0,488	0,744	300	14

With the specified ain Table 2 data, there are calculated the dependencies of the yield strength as a function of the generalized indicator (Re = f (Cekv)) and the elongation as a function of the generalized indicator (A = f (Cekv)). For specified models a test is carried out for validity of R with F – Fisher criterion. In Table 3 and Table 4 are shown derived dependencies for the yield strength (Re) and the elongation (A) with the respective coefficients of multiple correlation. In Fig.2 and Fig.3 a graphic interpretation of the models in Table 3 and Table 4 is shown.

Table 3. Models of the yield point according to the generalized indicator of the chemical composition

№	Models for the yield strength according to the generalized indicator of the chemical composition	Coefficient of multiple correlation
1	$Re=-202887x^4+457383x^3-368535x^2+126423x-14997$	$R^2=0,6496$
2	$Re=-157,69x^3+354,49x^2+531,28x+87,944$	$R^2=0,5095$
3	$Re= -790,13x^3+2496,2x^2-1565,1x+736,54$	$R^2=0,5352$
4	$Re= 9087,5x^3-12685x^2+6513,7x-584,03$	$R^2=0,6271$
5	$Re=-588,66x^3+1908,5x^2-1098,1x+572,44$	$R^2=0,6184$
6	$Re=-3397,9x^3+8292,7x^2-4721,7x+1358,8$	$R^2=0,5704$
7	$Re=-534,37x^3+1847,2x^2-1169,1x+599,78$	$R^2=0,6049$

Table 4. Models of elongation depending on the generalized indicator of the chemical composition

№	Models for elongation depending on the generalized indicator of the chemical composition	Coefficient of multiple correlation
1	$A=-19324x^5+55254x^4-61964x^3+34048x^2-9170,2x+981,37$	$R^2=0,5816$
2	$A=11,437x^2-36,248x+34,913$	$R^2=0,5546$
3	$A=11,427x^2-35,658x+33,646$	$R^2=0,5853$
4	$A=59,482x^2-87,718x+41,375$	$R^2=0,5415$
5	$A=11,11x^2-35,582x+35,031$	$R^2=0,6144$
6	$A=596,17x^4-1751,2x^3+1828,6x^2-810,06x+140,97$	$R^2=0,5228$
7	$A=10,135x^2-35,018x+36,246$	$R^2=0,6054$

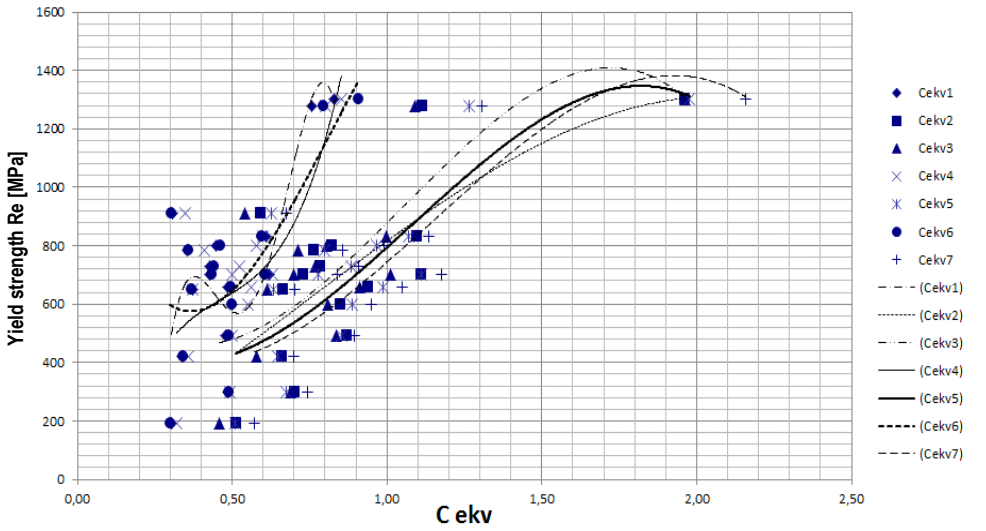


Fig.2. Graphic interpretation of the approximation of the experimental data for the yield strength.

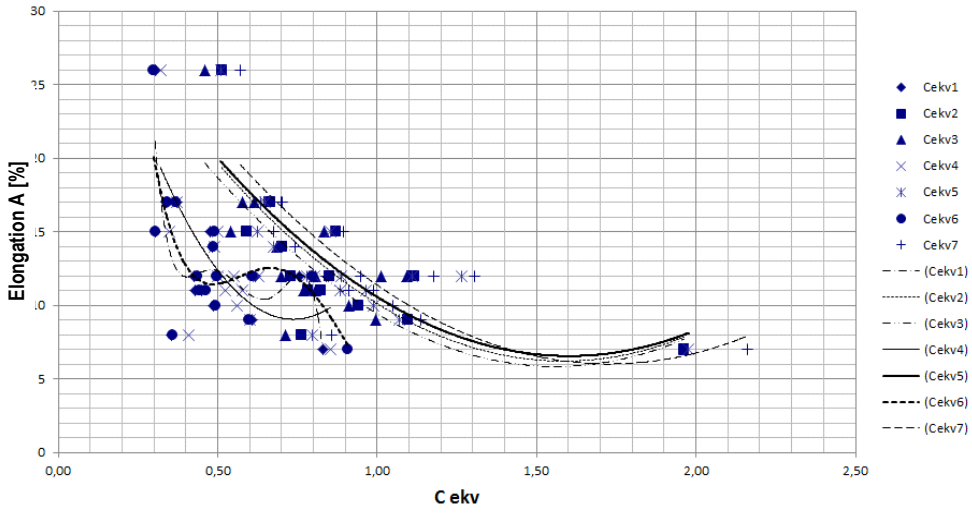


Fig.3. Graphic interpretation of the approximation of the experimental data for elongation.

Table 5. Efficient solutions for the vector of the generalized indicator of the chemical composition

Solution # 1	Solution # 2	Solution # 3
$\mathbf{B} := \begin{pmatrix} 0.48 \\ 0.85 \\ 0.8 \\ 0.52 \\ 0.9 \\ 0.5 \\ 0.9 \end{pmatrix}$	$\mathbf{B} := \begin{pmatrix} 0.48 \\ 0.85 \\ 0.8 \\ 0.52 \\ 0.9 \\ 0.5 \\ 0.95 \end{pmatrix}$	$\mathbf{B} := \begin{pmatrix} 0.48 \\ 0.85 \\ 0.8 \\ 0.505 \\ 0.9 \\ 0.5 \\ 0.9 \end{pmatrix}$

2.4. Results for optimal composition

As a result of Pareto optimization there have been received three effective solutions, which are given in Table 5. They are related to systems of linear equations, from the solution of which the optimal composition of the alloy is obtained. Table 6 presents the obtained Pareto solutions.

Table 6. Optimal chemical composition of effective solutions.

Chemical composition	Solution # 1	Solution # 2	Solution # 3
Carbon	$\mathbf{X} = \begin{pmatrix} 0.337 \\ 0 \\ 0.491 \\ 1.136 \\ 1.358 \\ 0.466 \\ 0.202 \end{pmatrix}$	$\mathbf{X} = \begin{pmatrix} 0.316 \\ 0.3 \\ 0.608 \\ 1.136 \\ 1.264 \\ 0.47 \\ 0.301 \end{pmatrix}$	$\mathbf{X} = \begin{pmatrix} 0.325 \\ 0 \\ 0.884 \\ 0.859 \\ 1.044 \\ 0.501 \\ 0.305 \end{pmatrix}$
Silicon			
Manganese			
Nickel			
Chromium			
Molybdenum			
Vanadium			

3. Testing the procedure

The idea of this section is to determine the influence of the accuracy of the models approximating the generalized indicator of the chemical composition over the final decision of the methodology for determining the optimal composition with a small database. For this purpose a priori are known the chemical composition and the relevant characteristics. The composition defined as a basic chemical composition fixes the different elements fixed as follows: C = 0.35%, Si = 1.4%; Mn = 0.52%; Mo = 0.53%; V = 0%, and the content of the other two elements Ni and Cr corresponds to the last 15-th row of Table 7 as follows:

Table 7. Chemical composition and properties of testing solutions.

Ni	Cr	Rm	Re	A	KCU	HB
1 [%]	1,5 [%]	1340,67 [MPa]	1237,84 [MPa]	16,56 [%]	888,81 [kJ/m^2]	244,91

The numeric verification can be performed with any combination in table 3.8, but the composition that was selected is: Ni = 1.0% and Cr = 1.5%, with a relatively optimal composition in terms of price/features.

According to the accepted method the solutions of the models are determined numerically rather than graphically and in this sense the differences between the preliminary composition and the calculatied one will be just from the accuracy of the regression model.

Table 8. Data about the generalized indicator of the chemical composition for the sample database.

№	Cekv1	Cekv2	Cekv3	Cekv4	Cekv5	Cekv6	Cekv7	A [%]
1	0,64	1,316	1,281	0,712	1,295	0,646	1,529	18,8
2	0,627	1,266	1,231	0,7	1,245	0,634	1,479	18,9
3	0,602	1,166	1,131	0,675	1,145	0,609	1,379	18,6
4	0,577	1,066	1,031	0,65	1,045	0,584	1,279	17,8
5	0,552	0,966	0,931	0,625	0,945	0,559	1,179	16,4
6	0,651	1,333	1,298	0,745	1,339	0,657	1,573	18,9
7	0,638	1,283	1,248	0,733	1,289	0,645	1,523	18,9
8	0,613	1,183	1,148	0,708	1,189	0,62	1,423	18,6
9	0,588	1,083	1,048	0,683	1,089	0,595	1,323	17,7
10	0,563	0,983	0,948	0,658	0,989	0,57	1,223	16,2
11	0,631	1,303	1,268	0,685	1,259	0,637	1,493	18,7
12	0,618	1,253	1,218	0,673	1,209	0,625	1,443	18,8
13	0,593	1,153	1,118	0,648	1,109	0,6	1,343	18,7
14	0,568	1,053	1,018	0,623	1,009	0,575	1,243	17,9
15	0,543	0,953	0,918	0,598	0,909	0,55	1,143	16,6

The results in Table 8 are obtained varying the two elements Ni and Cr in the basic composition. They are the basis to derive the regression equations described in Table 9.

Table 9. Models of elongation depending on the generalized indicator of the chemical composition with the respective statistical evaluation

A1	R = 0,8947	$y=101808191,9x^6 - 369200733,1x^5+557523417,02x^4-448734855,04x^3 + 203030392,74x^2 - 48960597,05x + +4916215,17$ - 16.6
A2	R = 0,9783	$y = 70452,31x^6 - 489376,30x^5 + 1413227,66x^4 - 2171664,05x^3 + +1872756,18x^2 - 859244,88\ x + 163865,66$ – 16.6
A3	R = 0,9783	$y = 70452,31x^6 - 474722,22x^5 + 1329672,46x^4 - 1981519,00x^3 + +1656891,09x^2 - 736998,03x + 136240,69$ - 16.6
A4	R = 0,5294	$y = -750841,85x^6 + 3588201,61x^5 - 6973575,95x^4 + 7093757,99x^3 - - 3998268,77x^2 + 1187038,65x - 145293,30$ - 16.6
A5	R = 0,8947	$y = 24855,52x^6 - 171993,65x^5 + 494551,74x^4 - 756281,29x^3 + +648615,20x^2 - 295740,36x + 56009,91$ - 16.6
A6	R = 0,8947	$y=101808191,92x^6 - 373018540,32x^5 - 569120593x^4 -462817656,61x^3 + 211575744,52x^2 - 51551610,36x + 5230260,18$ – 16.6
A7	R = 0,8947	$y = 24855,52x^6 - 206791,37x^5 + 715509,67x^4 - 1317819,04x^3 + +1362520,38x^2 - 749736,21x + 171525,98$ - 16.6

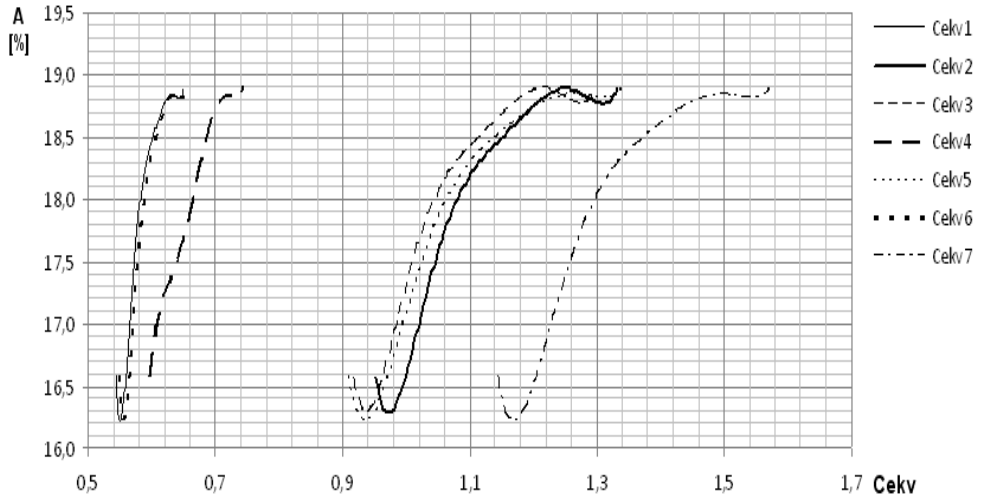


Fig. 4.9. Graphic interpretation of the elongation patterns for different generalized indicators of the chemical composition

The comparison between different elements from the chemical composition related to the initial baseline and the experimental numerical solution is presented in Table 10 . The test in this section was carried out at the Department of "Mathematics and Informatics" with the tool "Mathematica". The peculiarity of this processing is characterized by a high accuracy of the calculations.

Table 10. Comparison of the chemical composition from the basic solution with the one from the numeric processing

Chemical element	Basic composition %	Composition from the numeric solution %
Carbon - C	0.35	0.34998405
Silicon - Si	1.40	1.40306231
Manganese - Mn	0.52	0.51317959
Nickel - Ni	1.00	1.01919001
Chromium - Cr	1.50	1.499990306
Molybdenum - Mo	0.53	0.527142593
Vanadium - V	0.00	- 0.000336753

All models in Table 9 are of the sixth degree and thus they have six roots.

Two of them are real and four are complex. The solutions from the smaller roots are in Table 4.15. the solution with the bigger root is: C = 0.5594%, Si = 1.001%, Mn = 0.96554 %, Ni = 0.5980 %, Cr = 0.9736%, Mo = 0.5658%, V = 1.205%.

If this second solution is done then martensite will not be bundled, and the large amount of molybdenum and vanadium are in undeserved abundance.

4. Conclusion

A numerical procedure for determination of the optimal composition of a steel alloy using a small-size database with information for composites and their properties in one and the same process. The procedure was tested for class of alloy steels applied in energy-engineering industry. An optimal composition of the steel is obtained, which guarantees properties: yield strength Re > 600MPa and elongation values of A > 11%. The results show an acceptable match between the pre-known data with those obtained from numerical calculations determined by the presented methodology. The accuracy of the results is controlled by the used data. A good match is obtained with good statistical estimates. This approach should be established as one possible application, due to the significant reduction of the a priori information necessary for designing the optimal composition.

Acknowledgements:

The authors are grateful to fund ‘Scientific Research’, which partially funded the study under contract DDVU 02/11.

References

1. Gorni A.A (2013.). *Steel forming and heat treating.* Brazil.
2. *Catalogue of steel grades of 'EMSS.* (2011). Ukraine: Donetsk region.
3. Koprinkova-Hristova P., N. Tontchev, S. Popova, (2011), Neural Networks Approach to Optimization of Steel Alloys Composition,
4. Engineering Applications of Neural Networks. 315 - 324.
5. Koprinkova-Hristova P., N.Tontchev, S.Popova, (2013), Two approaches to multi-criteria optimization of steel alloys for crankshafts production, Int. J. Reasoning – based Intelligent Systems, Vol. 5, № 2, 2013, pp. 96 – 103.
6. Tontchev N., L. Kirilov, (2007). Two Approaches for Solving Multiple Criteria Decision Making MCDM Problems with an Illustrative Example. *Problems of Engineering, Cybernetics and Robotics, 58,* , 53-63,.
7. Tontchev N., N. Hristov, Numerical Procedure to Determine the Optimal Composition of the Steel a Small Volume Database with the Same Treatment, International virtual journal for science, technics and innovations for the industry, Machines, Technologies, Materials, VII, Issues 11, pp. 54 – 57, 2013.
8. Tontchev N., Y. Kalev, (2013), Robust Bi-criteria Approach to Optimize the Composition and Properties of Alloy Steels, Proceeding University of Transport, Vol 4.

Method of Authentication on Based Password Generators

A.F. Verlan, M.M. Karimov, K.A. Tashev, A.T. Imamaliev

Tashkent University of Information Technology, Uzbekistan,
dr.mmkarimov@rambler.ru, k.akhmatovich@gmail.com

Abstract. In this article the method of use of generators of pseudorandom numbers is offered at process of the authentication of the turned subject for protection of resources against unauthorized access in information and communication systems

Key words: pseudorandomness, cryptography, authentication, one time passwords, subject.

1. Introduction

Nowadays all of we knew results of different threats and attacks for unauthorized access to resources of information and communication systems.Therefore, development of measures for information and communication systems from unauthorized access is actual.Generally at protection of information and communication systems of users identification and authentication of users is used. Especially application of reliable, effective and one-time passwords at the authentication reduces probability of unauthorized use of system from malefactors. When forming such one-time passwords gives higher effect use of generators of pseudorandom numbers.

The following methods of application of odnoraovy passwords exist for authentication of users [1]:

- use of vrmenny values on the basis of uniform temporary system;
- use of the general lists of random numbers and the reliable mechanism of synchronization for the legal user and checking;
- use of generators of pseudorandom numbers with the identical initial values, being the general for the user and checked.

In these methods there are following shortcomings:

- complexity of synchronization of time for both parties;
- generated pseudorandom number it is valid in a certain period, and at this time the malefactor can use it.

2. Problems of authentication process

Taking into account the given shortcomings this way is offered. in an offered method authentication of users is realized through odnarozovy passwords with use of generators of pseudorandom numbers. The general view of system of authentication is given in figure - 1. (Fig. 1).

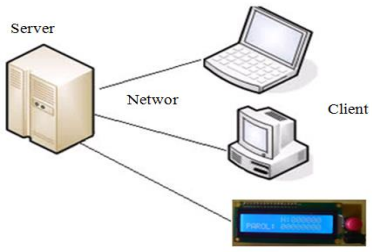


Fig.1. General view of system.

Server doing following functions:
- registration of the user via the hardware;
- the authentication of users via the hardware and saving of data in base.

User of the hardware:
- registration in the server;
- withdrawal of the password and number, which formed in a hardware window;
- Saving of the data in the server;

Generators of pseudorandom numbers are developed on the basis of the following methods:

1 . linear congruential.
2 . Multiplicative congruential.
3 . Generators on the basis of shift registers
4 . Generators on the basis of unilateral functions, algorithms of enciphering and hashing.
5 . Generators on the basis of problems of the theory of numbers.

it is possible to consider in detail some methods from them.

3. Method of generate pseudorandom numbers

Linear and multiplicative congruential generators. Linear Congruential Generator **(LCG)** yields a sequence of randomized numbers calculated with a linear equation. The method represents one of the oldest and best-known pseudorandom number generator algorithms. The theory behind them is easy to understand, and they are easily implemented and fast, especially on computer hardware which can provide modulo arithmetic by storage-bit truncation.

The generator is defined by the recurrence relation:

$$x_{i+1} = (ax_i + c) mod\ N$$

where X is the sequence of pseudorandom values, and

N, 0 <N — the "modulus"
a, 0 < a < N — the "multiplier"
c, 0 <= c < N — the "increment"
$X_0, 0 \leq X_0 < m$ — the "seed" or "start value"

are integer constants that specify the generator. If $c = 0$, the generator is often called a *multiplicative congruential generator* (MCG)

$$x_{i+1} = (ax_i) mod N, i = 0,1,2,3, \dots;$$

or Lehmer RNG. If $c \neq 0$, the method is called a *mixed congruential generator* [1].

Generators on based shift registers. Currently predlozhischie and the promising realizirovanie encrypt algorithm based on smooth shift registers or frankly form a linear feedback shift register. Such registers feedback shift register can say Fibonacci or Galois register. You can show two mend this type to use encrypted uninterrupted algorithms:

- Generators based on feedback shift registers gives the implementation of sequences of random static characteristics;
- Easy to analyze characteristics of generators based on shift registers.

Delhi itself feedback shift registers are divided into a linear feedback shift register and linear feedback shift registers.

General view of a linear feedback shift register is shown in Figure - 2.

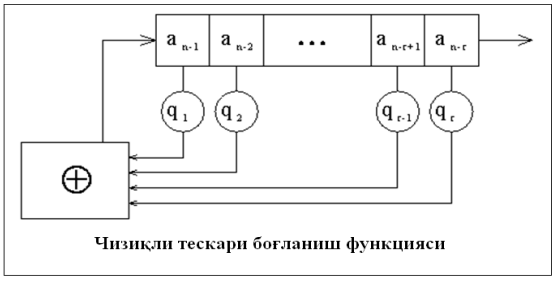


Fig.2. General view of the scheme of a linear feedback shift register

4. Process of generator's designing

Generators based on the shift register consists of two parts: The first part is the shift register, and the second part of the function on the feedback shift register. Implementation of algorithms based on shift registers that apply advanced software or hardware software aspect. To ensure and accelerate the creation of hardware and usability of technical devices equalized by microprocessors register with shift registers. For Intel processors running on 64- bit registers in the software registers size length equal to 64, and for the purpose of consent quadratic numbers, which are based on the algorithm of shift registers[4].

The analysis of methods of generation of pseudorandom numbers shows that for the authentication of subjects in information and communication systems the most effective is the linear and multiplicative congruent method when developing the hardware device of generation of one-time passwords. Because, this method distinguishes from others simplicity of hardware realization.

In this scientific article the way of development of the hardware device of generation of pseudorandom numbers on the basis of a linear and multiplicative congruent method is offered.

For development of the hardware device of generation of pseudorandom numbers it is recommended to use an ISIS package of the Proteus program and it has to contain the following elements:

- LCD;
- microcontroller (Atmel Atmega 16);
- Resistor;
- Button;
- Condenser.

The scheme of the hardware device of the generator of pseudorandom numbers in the Proteus program to look as follows (fig.3).

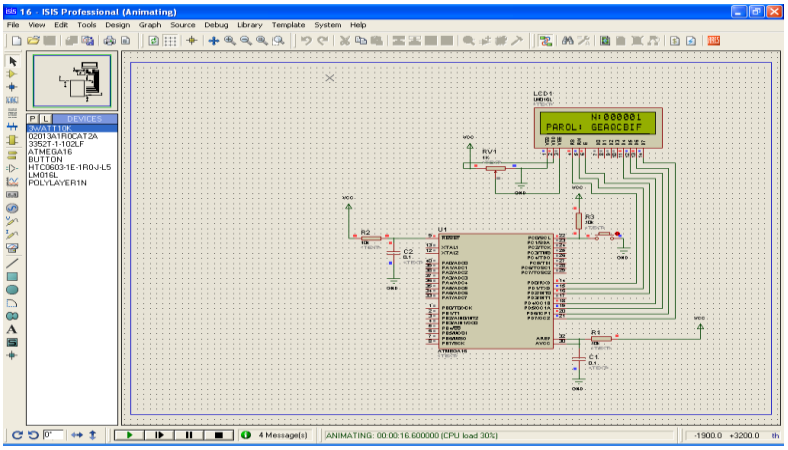


Fig.3. The scheme of the hardware device of the generator of pseudorandom numbers

At the following stage the developed program in language "C" for generation of sequences on the basis of a linear and miltiplicativecongruent method it is necessary translates in the Assembler language which understands a microchip by means of the CodeVision AVR [5] program.

At the following stage by means of a programmator it is carried out flashing programs for memory of a microchip.

If there are no changes and mistakes that it is considered successfully executed flashing programs with the HEX format on the Atmega16 [5] microcontroller.

Fig.4. View of hardware.

5. Conclusion

In conclusion can to speak, that one more of features of this hardware device, it is each generated symbol it is created on separate microprogramming which gives the chance to increase firmness of generated pseudorandom numbers.

References

1. S.K.Ganiev, M.M. Karimov, K.A. Tashev. «Information security.Information and communication systems security». Tashkent. «Aloqachi» 2008y, P.-378.
2. S.S.Kasimov. Information technologies. Tutorial.-T.: «Aloqachi», 2006.
3. S.K. Ganiev, T.A.Kuchkarov, X.N.Gulomova. «Circuit technology».-T.: «Aloqachi». 2009 y., P-148.
4. PetrovA.A. «Computersecurity. Cryptographic methods of protection », M.: DMK, 2000y. P.-448.
5. Evstifeev A.V. Microcontrollers AVR of famile Tiny and Mega firm«ATMEL».-M.: "Dodeka-XXI", 2004y.P.-560.
6. GanievS.K., KarimovM.M. «An integrated system of information security in corporate networks»./Theses of reports of the International SCP «Innovation 2001».-Tashkent, 2001.-P.97-98.
7. GanievS.K., KarimovM.M. «Combined method of crypting to protect information systems»./Theses of reports of the RepublicSCP«Technical disciplines and global issues XXI century ».-Tashkent, 2001.-P.12-13.

An Algorithm for Computing of Critical Loads for Heavy Metals for Aquatic Ecosystems

Sonya Damyanova

LTU, 1756 Sofia, Bulgaria
sonya_damyanova @abv.bg

Abstract. The paper suggests a calculation model for computing critical loads for lead and cadmium for aquatic ecosystems. The general model is combined with some measured water characteristics to estimate a real state of a water body. All steps of calculation and data source are descripted in details. An example for a calculation of critical loads for a Chiprovci River ecosystem is presented as well.

Keywords. Critical Loads, Heavy Metals, Aquatic Ecosystem, Atmospheric Deposition.

1. Introduction

Critical load approach is an integral way to evaluate ecosystem sustainability against an anthropogenic impact in a long-term period. The main concept is SSMB (Steady State Mass Balance) model that is based on the equilibrium of input and output mass fluxes go into and out of a considered ecosystem [2, 3]. That way critical loads take into account all individual features of an ecosystem and describe its sensitivity against pollutants.

That model is widely used for a risk assessment of heavy metal pollution impact in a frame of International Cooperative Program (ICP) Forest and a Center of Effects working group at Modelling and Mapping Program. Numerous of articles [4, 6] report date of critical loads for forest ecosystems but information in respect of aquatic ecosystems is limited [8].

UBA Manual for calculation of critical loads for lead, cadmium and mercury [7] describes a model based on common water characteristics such as pH, dissolved organic carbon etc. Transfer functions are used to compute different metal concentrations. This paper suggests a model that uses measured parameters connected to metal sediment content and amount of suspended particulate matter, which define critical metal concentration in surface water. In this way one can estimate a real state of ecosystem and take measures to protect its normal functioning in a future.

Input data

Critical loads combine metal fluxes flowing throughout the whole ecosystem in a respect of its safety and normal functioning. One of the metal pathway is uptake by vegetation and the other one is element outflow from the whole catchment. So data on catchment area, species composition, annual yield of biomass and Pb, Cd contents in the harvestable part of vegetation is needed for estimation of metal remove by biomass. Metal output fluxes from the watershed depend on outflow flux of water from the catchment, concentrations of suspended particles in surface water and Pb and Cd contents in the suspended particles.

2. Algorithm steps

To estimate heavy metal removal by net uptake have to calculate element release from the whole catchment. The metal output by biomass can be calculated following 1st to 4th steps. Equations 5th to 7th are used to assess a critical metal output by lateral outflow water flux. The last step summarize the both metal fluxes forming critical load as an acceptable total load of anthropogenic heavy metal input.

1. Calculation of annual yield of harvestable part of vegetation using following equation:

$$Ya = Am \,.\, d \quad (1)$$

where: Ya is an annual yield of harvestable part of vegetation as a dry weight, $kg.ha^{-1}.yr^{-1}$;
Am – a volume annual increment, $m^3.ha^{-1}.yr^{-1}$;
d – density of a dry wood, $kg.m^{-3}$.

For a mix stand it is recommended to divide vegetation to deciduous and coniferous species and make the calculation separately. Data on wood density can be found in a literature.

2. Calculation of the annual net yield of biomass in the catchment according to:

$$Yha = Ya \,.\, P \quad (2)$$

where: Yha is an annual net yield of biomass in the catchment, $kg.ha^{-1}.yr^{-1}$;
P – species composition part in the catchment, $kg.ha^{-1}.yr^{-1}$.

Data for the species composition can be found in a land cover database. Data on the yield for forest can be obtained directly from the related national statistics.

3. Calculation of the annual metal net uptake of the harvestable part of plants in the catchment is given by:

$$Mu = Yha \,.\, [M]ha \quad (3)$$

where:
Mu is a metal net uptake in harvestable part of plant under critical load conditions, $g.ha^{-1}.yr^{-1}$;
[M]ha – the metal content in harvestable plant part in a dry wood, $g.kg^{-1}$.

The metal content of Pb varies between 0.0005 – 0.010 $g.kg^{-1}$ for Central Europe. The values of Cd for deciduous forest are: 0.00005 – 0.0005 $g.kg^{-1}$ and for coniferous: 0.0001 – 0.0005 $g.kg^{-1}$ respectively [3]. Data for the metal content should be taken from unpolluted areas as average values from all part of trees including a bark. As the leaves and needles remain in the catchment, they have to be removed from the calculation. Values of 0.006 $g.kg^{-1}$ and 0.007 $g.kg^{-1}$ for Pb content of Bulgarian deciduous forests for coniferous was obtained. For Cd the related values are: 0.0003 and 0.0004 $g.kg^{-1}$.

4. Calculation of the total annual metal net uptake of the harvestable part of plants in the catchment according to:

$$Mtot = Mdec + Mcon \quad (4)$$

where: Mtot is the total annual metal net uptake of the harvestable part of plants in the catchment, $g.ha^{-1}.yr^{-1}$;
Mdec – the metal deciduous net uptake of biomass, $g.ha^{-1}.yr^{-1}$;
Mcon – the metal coniferous net uptake of biomass, $g.ha^{-1}.yr^{-1}$.

To calculate critical metal outflow from the whole catchment, following steps should be done.

5. Calculation of the critical metal content in suspended particles using an equation:

$$[M]spм(crit) = [SPM]sw \,.\, [M]spм \quad (5)$$

where: [M]spм(crit) is the critical metal content in suspended particles, $mg.dm^{-3}$;
[SPM]sw – a concentration of suspended particulate matter, $g.dm^{-3}$;
[M]spм – a metal concentration in water sediments, $mg.kg^{-1}$.

6. Calculation of the critical total concentration of heavy metal in surface water following an equation:

$$[M]tot,sw(crit) = [M]dis(crit) + [M]spм(crit) \quad (6)$$

where:
[M]tot,sw(crit) is the critical total concentration of heavy metal in surface water, $mg.m^{-3}$
[M]dis(crit) – the critical dissolved metal concentration in surface water, $mg.m^{-3}$.

Recommended values for critical concentration of dissolved Pb in surface water is 11 [1] and: 0.38 $mg.m^{-3}$ respectively [5]. These limits are related to ecotoxicological effects for aquatic ecosystems. The values for critical metal concentrations were obtained from laboratory tests for chronic toxicity of different aquatic organisms based on dose-response functions.

7. Calculation of the critical lateral metal outflow flux from the catchment is given by:

$$Mlo(crit) = 10 \,.\, Qlo \,.\, [M]tot,sw(crit) \quad (7)$$

where: Mlo(crit) is the critical lateral metal outflow flux from the catchment, $g.ha^{-1}.yr^{-1}$;
Qlo – the lateral water outflow from the watershed, $m.yr^{-1}$.

It is necessary to use a transfer factor 10 (in $g.mg^{-1}.m^2.ha^{-1}$) to convert units from $mg.m^{-2}.yr^{-1}$ to $g.ha^{-1}.yr^{-1}$.

8. Calculation of the critical loads for heavy metals according to:

$$CL(M) = Mtot + Mlo(crit) \quad (8)$$

where: CL(M) is a critical load for heavy metal, $g.ha^{-1}.yr^{-1}$

3. An example

Critical loads for Pb and Cd for the ecosystem of Chiprovci River were calculated following the algorithm described above. Lead net uptake in harvestable part of plants is shown on the Figure 1. The investigated ecosystem consists of 31% coniferous (column B) and 69% deciduous species (column I). The coniferous volume increment (Fig. 1, column C) is twice as large as the broad-leaved one (Fig. 1, column J). In the contrary, the deciduous density (Fig. 1, column L) is higher than the coniferous one (Fig. 1, column D), so the annual yield of biomass is not so much different. Taking into account species composition in the ecosystem, it is obvious that the net deciduous growth (1353 $kg.ha^{-1}.yr^{-1}$) is about 50% higher than the coniferous net yield (816 $kg.ha^{-1}.yr^{-1}$). The Pb content in harvestable part of the broad-leaved forest (0.006 $g.kg^{-1}$) is lower than the one in the coniferous trees (0.007 $g.kg^{-1}$) because the coniferous species hold foliage during the whole year. Since the ecosystem contains mainly deciduous forest, the net Pb uptake by biomass of the broad-leaved trees (Fig. 1, column O, 8.12 $g.ha^{-1}.yr^{-1}$) is higher than the net coniferous one (Fig. 1, column H, 5.71 $g.ha^{-1}.yr^{-1}$). Therefore the Pb uptake from the whole ecosystem in harvestable part of plants is 13.82 $g.ha^{-1}.yr^{-1}$ as it can be seen in the Fig. 1, column P.

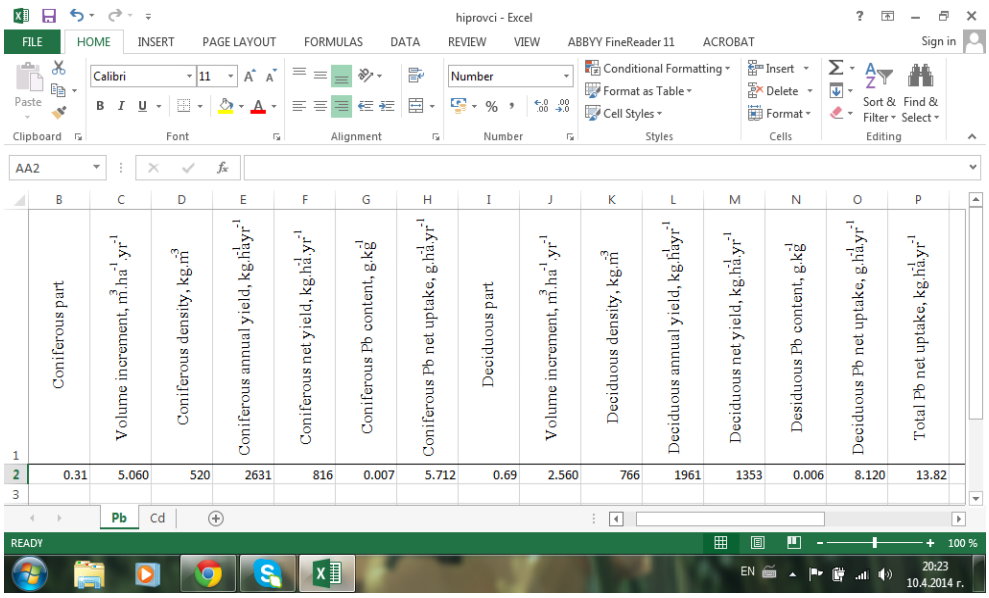


Fig. 1. Calculation of Pb net uptake by biomass.

Figure 2 displays calculation steps for estimation of critical lateral lead outflow flux from the same watershed of Chiprovci River. It is shown a critical load of lead in it as well.

The annual measured suspended particles in the surface water of the river and lead sediment concentration respectively are shown in the Fig. 2 columns S and T. Since it is difficult to obtain metal concentration that is adsorbed at suspended matter, we accepted that Pb content is the same as one in the sediments. Hence column U in Fig. 2 shows the lead concentration in the suspended particles. Calculated total critical Pb concentration in surface water of Chiprovci River (Fig. 2, column W) is higher than the critical dissolved Pb

concentration (Fig. 2, column V). That big difference between total and dissolved lead concentrations is due to a long term mining activity in the watershed of the river. In case of lack of sediment data it is possible to use dissolved metal concentration as a total one, especially for unpolluted aquatic ecosystem. Fig. 2, column X shows the calculated critical lead outflow (9.18 $g.ha^{-1}.yr^{-1}$). The last column Y in the same figure calculates critical load for Pb for the surface water of Chiprovci ecosystem (23 $g.ha^{-1}.yr^{-1}$).

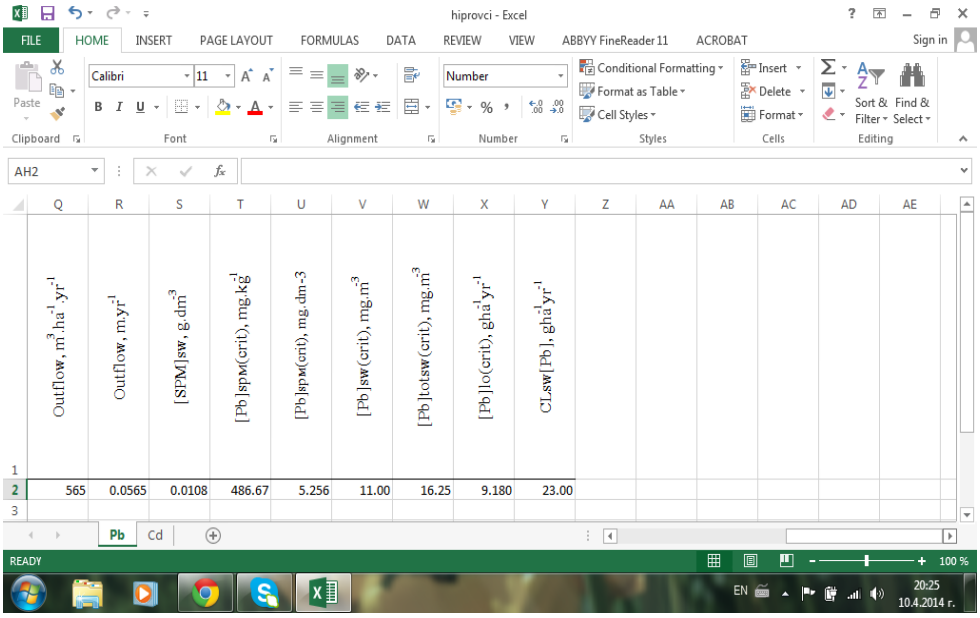


Fig. 2. Calculation of Pb net outflow.

Following two figures (Fig. 3 and 4) display the same calculation steps that were described above but for cadmium. Nevertheless, species composition and annual yield are the same for Chiprovci ecosystem, calculated Cd uptake by biomass (0.41 $g.ha^{-1}.yr^{-1}$) is 20 times as lower as Pb one (Fig. 3, column P).

As it can be seen in the Fig. 4, column T, the cadmium sediment content (44.36 $mg.kg^{-1}$) is about 10 times less than the lead one. Therefore the critical Cd concentration in suspended particles is less 10 times as well (Fig. 4, column U).

The calculated total critical Cd concentration (Fig. 4, column W) is twice higher than critical dissolved one (Fig. 4., column V). Therefore it is recommended to pay attention on polluted ecosystems and measure actual concentrations regularly. Otherwise using only recommended critical dissolved concentration value, the calculated result for metal outflow from ecosystem will be lower an critical load will not describe the real situation.

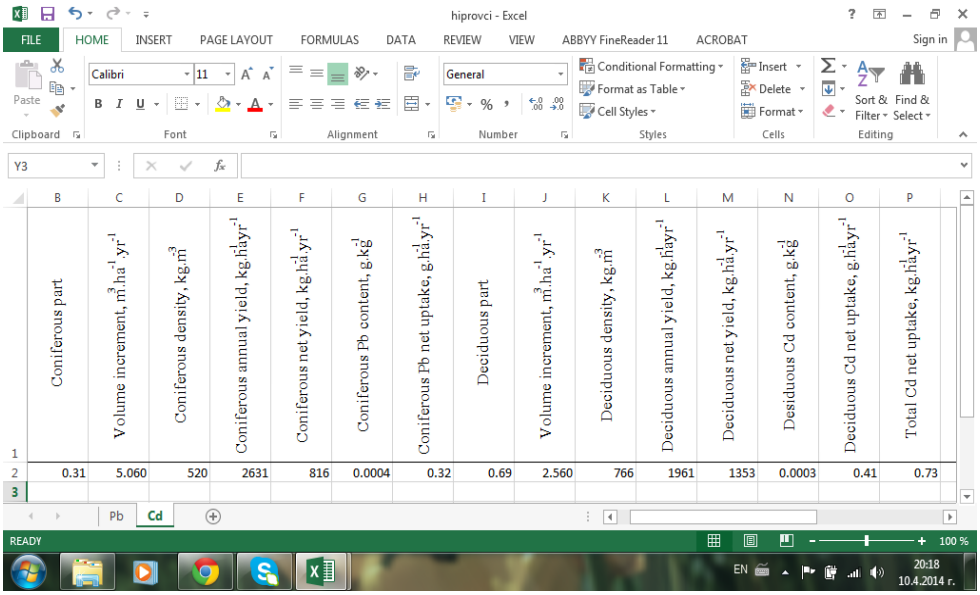


Fig. 3. Calculation of Cd net uptake by biomass.

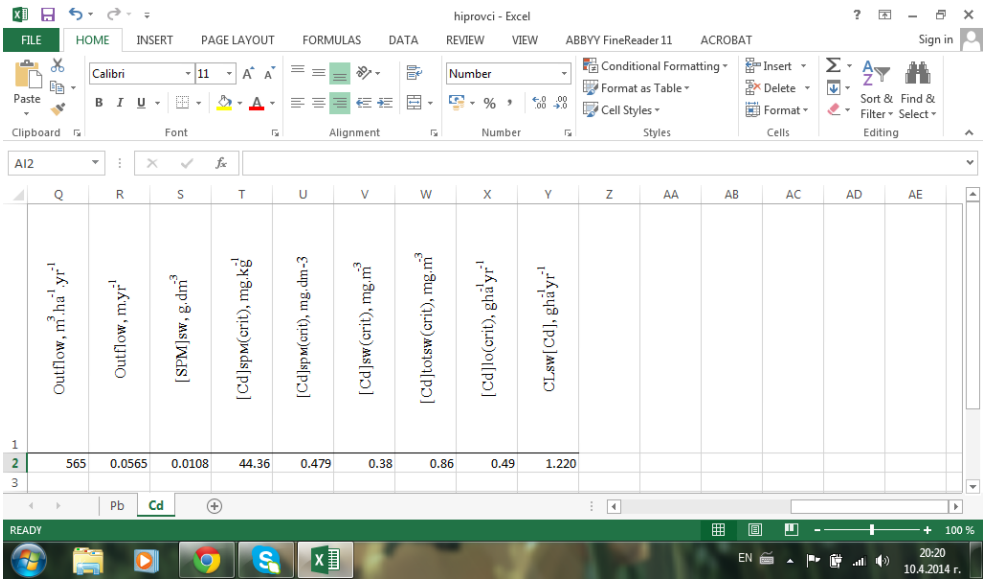


Fig. 4. Calculation of Cd net outflow.

4. Conclusion

Calculation of critical loads for lead and cadmium is based on a simple Steady State Mass Balance approach for input and output metal fluxes flowing throughout a considered ecosystem. The algorithm is similar in its first part to that for calculation of critical loads for heavy metals for forest ecosystems. An acceptation for equal metal content in sediments and suspended particulate matter is a good agreement for the calculation procedure. In a case of polluted sediments the difference between dissolved and total critical metal concentrations will be significant. To estimate a real state and sensitivity of an ecosystem is better to use measured values of parameters. The calculated critical loads for Ciprovci River ecosystem showed a significant higher sensitivity of the ecosystem to cadmium compared to lead.

References

1. Crommentuijn T., Polder M.D., Van de Plassche E.J. Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account. National Institute of Public Health and the Environment, Bilthoven, The Netherlands. Report 601501 001, 157 p, 2007.
2. De Vries, W., Bakker, D.J., Sverdrup, H.U. Manual for calculating critical loads of heavy metals for aquatic ecosystems. Guidelines for critical limits, calculation methods and input data. DLO Winand Staring Centre, Report 165, Wageningen, The Netherlands, 91 p, 1998.
3. De Vries W., Schütze G., Lofts S., Tipping E., Meili M., Römkens, P.F.A.M., Groenenberg J.E. Calculation of critical loads for cadmium, lead and mercury. Background document to a Mapping Manual on Critical Loads of cadmium, lead and mercury. Wageningen, Alterra, Alterra Report 1104: 143 p, 2005.
4. Downing R, Hettelingh J-P, De Smet PAM. Calculation and mapping of critical loads in Europe: Status Report 1993. CCE/ RIVM, Rep. 259101003. Bilthoven, 1993.
5. European Commission. Risk Assessment Report on Cadmium metal and Cadmium oxide. ECB web-site: http://ecb.jrc.it, 2003.
6. Hettelingh J.P., Posch M., Slootweg J. Critical load, dynamic modelling and impact assessment in Europe: CCE Status Report 2008, Coordination Centre for Effects, www.pbl.nl/cce, 2008.
7. UBA. Modelling and Mapping Manual: Manual on methodologies and criteria for modelling and mapping critical loads & levels and air pollution effects, risks and trends, Federal Environmental Agency (Umweltbundesamt) Berlin, UBA-Texte 52/04, 2004.
8. Slootweg J., Hettelingh J.-P., Posch M., Schütze G., Sprander T., De Vries W., Reinds G., J., Van't Zelfde M., Dutchak S., Ilyin I. *European critical loads of cadmium, lead and mercury and their exceedances*. Water Air Soil Pollution 7; 371-377, 2007.

A Model for Calculation of Critical Loads for Lead and Cadmium for Surface Waters

Sonya Damyanova

LTU, 1756 Sofia, Bulgaria
sonya_damyanova@abv.bg

Abstract. The paper offers a simple and useful way to estimate surface waters sensitivity to heavy metal pollutants. Critical load approach based on Steady-State Mass Balance model is used. Critical total metal concentration in surface water is calculated using transfer functions. An example for stream water is shown. Calculated critical load value for cadmium is much lower than lead one. Therefore a risk of damage for aquatic organisms from Cd is significant higher than Pb one.

Keywords. Critical loads, Heavy metals, Surface waters, Critical metal concentration.

1. Introduction

Critical load approach is an appropriate one for an estimation of ecosystem sensitivity against anthropogenic impact. According the definition of Nilsson and Grennfelt [6] critical loads are these acceptable thresholds of pollutants that do not cause significant harm of an ecosystem element or ecosystem functioning. That approach has been successfully applied to solve problems with acidification and eutrophication of surface waters in Europe under the Convention on Long-Range Trans-boundary Air Pollution (CLRTAP) [1].

Calculation of critical loads is based on a simple Steady-State Mass Balance model for input and output fluxes from a considered ecosystem. When an object of investigation is surface waters the main problem is how to calculate critical total metal concentration of water body. Data on critical dissolved metal concentration are usually obtained from laboratory toxicity tests for different organisms. Lots of aquatic organisms could be harmed by adsorption of toxic particles on their bodies in permanent contact with water or when they swallow that matter. Therefore it is important to estimate not only critical dissolved element content but critical amount adsorbed on suspended particulate matter as well.

Tipping [7, 8] suggests a computing model for calculation of critical total metal concentration for surface waters. Input data are common water characteristics such as pH, dissolved organic carbon and organic matter content in sediments. The model is based on transfer functions to calculate critical metal bond to particulate matter from critical free metal ion concentration.

2. Method for calculation of critical loads for heavy metals

A simple steady-state mass balance approach is used to calculate critical loads of heavy metals. The method is based on an assumption of equilibrium of all relevant input and output metal fluxes from the studied ecosystem [3, 9]. The model implies that the critical load equals the sum of acceptable net outputs by harvest and outflow from the catchment, minus inputs by natural weathering release plus net retention. Usually there is a lack of

information about net retention for many of still waters, so that it is possible to ignore that component of an equation. When stream water is an object of investigation, there is no need a net retention to be included as well [4]. The heavy metal net release is negligible and its estimation includes high uncertainties. In that way there is no need metal weathering to be considered [4]. As a result of all considerations mentioned above, the critical load of heavy metal can be calculated as a sum of output metal fluxes from the catchment by biomass and outflow:

$$CL(M) = Mu + Mlo(crit) \quad (1)$$

where: CL(M) is a critical load of a heavy metal, $g.ha^{-1}.yr^{-1}$;
Mu metal net uptake in harvestable parts of plants under critical load conditions, $g.ha^{-1}.yr^{-1}$;
Mlo(crit) – critical metal outflow from the catchment, $g.ha^{-1}.yr^{-1}$.

Metal net uptake by biomass (Mu) has to be calculated by multiplying the annual yield with the metal content of the harvestable parts of the trees, including metals adsorbed on the surfaces, as follows:

$$Mu = Yha \, . \, [M] \quad (2)$$

where:

Yha is a yield of harvestable biomass (dry weight), $kg.ha^{-1}.yr^{-1}$;
[M]ha – metal content of the harvestable part, $g.kg^{-1}$.

Data for the metal content of harvestable biomass should be taken from analysis of trees growing in unpolluted areas. Calculated fluxes can be therefore considered as tolerable. When data for different part of tree are available, the weighted mean concentration should be used. For forest ecosystems only the net increment (excluded leaves, needles) should be taken.

Critical metal outflow from the catchment (Mlo(crit)) has to be calculated according to the following equation:

$$Mlo(crit) = clo \, . \, Qlo \, . \, [M]tot,sw(crit) \quad (3)$$

where: Qlo is outflow flux of water from the catchment, $m.yr^{-1}$;
[M]tot,sw(crit) – critical total concentration of heavy metal in the surface water, $mg.m^{-3}$;
clo – 10, factor for appropriate conversion of units from $mg.m^{-2}.yr^{-1}$ to $g.ha^{-1}.yr^{-1}$.

The critical total metal concentration in surface water has to be calculated as a sum of critical dissolved metal concentration and critical concentration of metal bond to suspended particles:

$$[M]tot,sw(crit) = [M]\,sw(crit) + [M]_{SPM(crit)} \, . \, [SPM]sw \quad (4)$$

where:
[M] sw(crit) is a critical dissolved concentration of a heavy metal in surface water, $mg.m^{-3}$;
$[M]_{SPM(crit)}$ – critical total content of a heavy metal in suspended particles, $mg.kg^{-1}$;
[SPM]sw – concentration of suspended particles in surface water, $kg.m^{-3}$.

Recommended values for critical dissolved metal concentrations are 0.38 $mg.m^{-3}$ for Cd [5, 10] and 11 $mg.m^{-3}$ for Pb [2]. Since the dissolved metal concentration depends on hydrochemistry of water body, more precisely approach is to measure data for pH, dissolved organic carbon and concentration of suspended particles. Moreover in UBA [9] it is considered to count water hardness because of positive relation with dissolved metal concentration. De Vries at al. [4] suggest value 5 $mg.m^{-3}$ for Pb and for Cd: 0.16 $mg.m^{-3}$, when H (hardness) < 100; 0.30 $mg.m^{-3}$, when 100 < H < 200 and 0.50 $mg.m^{-3}$, when H > 200 mg $CaCO_3.dm^{-3}$.

3. Calculation of critical total Cd and Pb concentrations for surface water

Calculation procedure includes 4 steps.

1. Calculation of the critical free ion concentration.
2. Calculation of the critical concentration of metal bond to suspended particles.
3. Calculation of the critical dissolved metal concentration according to the hardness.
4. Calculation of the critical total metal concentration.

1. In computing model WAM6 [7, 8], Tipping suggests to estimate a critical free ion concentration depending on the pH, dissolved organic carbon (DOC) and pressure of CO_2 according to following equation:

$$\log[M]free,crit = A \cdot [DOC] + B \cdot p\,CO_2 + C \quad (5)$$

where:

A, B and C are regression coefficients;

p CO_2 – a multiple of the atmospheric p CO_2;

[DOC] – a concentration of dissolved organic carbon, $mg.dm^{-3}$.

For example for the catchment of Chiprovci River following mean annual values for: pH = 6.89, [DOC] = 0.6 $mg.dm^{-3}$ are measured. Thus the critical free ion concentration for cadmium equals:

$$\log[Cd]free,crit = (-0.0339 \cdot 0.6) + (0.186 \cdot 4) - 8.91 = -8.18$$

The same equation has to be used for calculation of critical free ion concentration for lead:

$$\log[Pb]free,crit = (-0.067 \cdot 0.6) + (0.0257 \cdot 4) - 9.18 = -9.14$$

2. The critical concentration of metal bond to suspended particles is calculated using regression coefficients connecting free metal activity and metal content bond to the solid phase as follows:

$$\log[M]_{SPM(crit)} = a + b \cdot \log[OM] + c \cdot pH + m \cdot \log[M]free \quad (6)$$

where: a, b, c and m are regression coefficients;

OM – organic matter content, % dry weight;

$[M]_{SPM(crit)}$ – reactive metal content in the solid phase, $mol.g^{-1}$.

For Chiprovci River, organic matter content is 20% and the reactive metal content in the solid phase for Cd equals:

$$\log[Cd]_{SPM(crit)} = -6.42 + (0.64 \cdot \log 20) + (0.45 \cdot 6.89) + (0.58 \cdot (-8.18)) = -7.23$$

and for Pb respectively:

$$\log[Pb]_{SPM(crit)} = -5.42 + (0.55 \cdot \log 20) + (0.7 \cdot 6.89) + (0.61 \cdot (-9.14)) = -5.45$$

To convert units from $mol.g^{-1}$ to $mg.kg^{-1}$ it is necessary to multiply corresponding concentration by atomic mass and 10^6.

$$[Cd]_{SPM(crit)} = 5.86.10^{-8} \cdot 112 \cdot 10^6 = 6.56 \text{ mg.kg}^{-1}$$
$$[Pb]_{SPM(crit)} = 3.49.10^{-6} \cdot 207 \cdot 10^6 = 723 \text{ mg.kg}^{-1}$$

3. Relationship between water components is used to calculate hardness (H) according to the equation:

$$H = A \cdot [DOC] + B \cdot pCO_2 + C \quad (7)$$

$$H = (0.31 \cdot 0.6) + (3.35 \cdot 4) - 0.12 = 13.46$$

According to De Vries at al. [4] recommended value for critical dissolved Cd concentration is 0.16 $mg.m^{-3}$. Critical dissolved Pb concentration is 5 $mg.m^{-3}$, independently on water hardness.

4. The total metal concentration at critical limit is calculated by equation 4:

$$[M]tot,sw(crit) = [M]\,sw(crit) + [M]_{SPM(crit)} \cdot [SPM]sw$$

Following results are obtained for Chiprovci river for Cd:

$$[Cd]tot,sw(crit) = 0.16 + (19 / 1000 \cdot 6.56) = 0.28 \text{ mg.m}^{-3}$$

and for Pb respectively:

$$[Pb]tot,sw(crit) = 5 + (19 / 1000 \cdot 723) = 19 \text{ mg.m}^{-3}$$

When calculated above values for critical total metal concentrations is put in equation (3), following results for critical metal outflow can be found:

$$Cdlo(crit) = 10 \cdot 0.0565 \cdot 0.28 = 0.16 \text{ g.ha}^{-1}.\text{yr}^{-1}$$

$$Pblo(crit) = 10 \cdot 0.0565 \cdot 19 = 10.73 \text{ g.ha}^{-1}.\text{yr}^{-1}$$

Critical loads can be calculated summarising metal outflow with metal net uptake by biomass according equation 1:

$$CL(Cd) = 0.73 + 0.16 = 0.89 \text{ g.ha}^{-1}.\text{yr}^{-1}$$

$$CL(Pb) = 13.82 + 10.73 = 24.55 \text{ g.ha}^{-1}.\text{yr}^{-1}$$

4. Conclusion

The described model offers a simple and useful way for calculation of critical loads for heavy metals for surface waters. A few water characteristics are needed for calculation of critical total metal concentration. There is a possibility to obtain relevant coefficients according specific water body features.

Calculated critical loads for Chiprovci River show a significant sensitivity of aquatic organisms to cadmium pollutant comparing to lead one. To assess a risk for ecosystem it is necessary to compare critical loads with annual metal deposition and to keep deposition level below critical load values. Critical load approach is an useful and practical tool for monitoring of waters and protection of aquatic ecosystems.

References

4. CLRTAP, Convention on Long-Range Transboundary Air Pollution, 1979.
5. Crommentuijn T., Polder M.D., Van de Plassche E.J, *Maximum permissible concentrations and negligible concentrations for metals, taking background concentrations into account*, National Institute of Public Health and the Environment, Bilthoven, The Netherlands. Report 601501 001, 157, 1997.
6. De Vries W., D.J. Bakker D. J., Sverdrup H. U, Manual for calculating critical loads of heavy metals for aquatic ecosystems, Guidelines for critical limits, calculation methods and input data. DLO Winand Staring Centre, Wageningen (The Netherlands). Report 165, 1998.
7. De Vries W., Schütze G., Lofts S., Tipping E., Meili M., Römkens P.F.A.M., Groenenberg J.E., Calculation of critical loads for cadmium, lead and mercury. Background document to a Mapping Manual on Critical Loads of cadmium, lead and mercury, Wageningen, Alterra, Alterra Report 1104: 143, 2005.
8. European Commission Risk Assessment, Report on Cadmium metal and Cadmium oxide, ECB, 2003, (http://www.ecb.jrc.it.).
9. Nilsson J., P. Grennfelt , *Critical loads for sulphur and nitrogen*, Report from a workshop held at Skokloster Sweden, 19-24, 1988.
10. Edward Tipping, *WHAM–A chemical equilibrium model and computer code for waters, sediments, and soils incorporating a discrete site / electrostatic model of ion-binding by humic substances*, Comput. Geosci. 20 (6), 973-1023, 1994.
11. Edward Tipping, *Humic ion-binding model VI: an improved description of the interaction of protons and metal ions with humic substances*, Aquatic Geochemistry 4 (1), 3-47, 1998.
12. UBA, Manual on methodologies and criteria for modelling and mapping critical loads & levels and air pollution effects, risks and trends, UNECE Convention on Longrange Transboundary Air Pollution, Federal Environmental Agency (Umweltbundesamt), Berlin: 254, 2004.
13. WHO, Guidelines for drinking water quality, Third edition, vol. 1- Recommendations, Geneva, 2004.