ADAPTIVE TRAFFIC SIGNAL CONTROL USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Daniela Koltovska Nečoska
Kristi Bombol
Sv. Kliment Ohridski University, Faculty of Technical Sciences
Department for Traffic and Transport
POB 99, 7 000 Bitola, Republic of Macedonia

Summary

A large number of intelligent systems have been developed to adapt to the flow changes as well as to be capable for effective management of the signal time changes.

For a long time it was believed that the systems responding to real time traffic would enable significant benefits. But, numerous limitations such as use of models with high level of detail precision have occurred. The conventional optimal control methods suffer from the so-called curse of dimensionality. The difficulties for optimal signal control and the importance of this problem lead to a great number of new researches. A possible solution for the above stated problems is discerned in the methods of the artificial intelligence. The artificial intelligent system reaches the same results as man does when performing cognitive tasks. These systems are characterized with the ability to accumulate and use knowledge, set the problem, learn, conclude, solve the problem, process and exchange knowledge.

The research presented in this paper proposes an adaptive signal control performed by a control agent that is able to adapt an optimal policy by learning from the environment. The goal to be achieved is to minimize the delays in the system.

The setting up of Q-learning algorithm and the first computation results of Q-learning application for adaptive traffic signal control will be presented. It is shown that the results obtained are in favor of adaptive signal control compared to the actuated signal control.

1. INTRODUCTION

The Intelligent Transport Systems (ITS) which are used by the advanced technologies in the transport system are widely dispersed approach as a solution for the traffic problems which the society faces with.

The aim to create ITS is improvement of the traffic quality by avoiding traffic congestions, saving travelling time, improving the safety and the comfort of the drivers and the passengers, improving the transport and the exchange of goods and services and improvement of the entire informative transparency. As a result of that, we receive a higher level of satisfaction for the user's needs and the entire prosperity of the surrounding area.

The traffic management and control in cities is just one of the areas where ITS can be applied.

Traffic signal control strategies are the most commonly used for managing the traffic flow for different types of intersections in the cities. At a signalized intersection they operate in one of the three different control modes: pre-timed control, semi actuated control and fully-actuated control (Wilshire et al., 1985).

During the last two decades significant efforts are made to develop efficient and practical real time strategies to control the traffic at intersections. Although the concept of these efforts is promising, the difficulties still exist. To predict the origin and destination traffic demand in real time, the inherent limits in modeling the complex traffic flows, as well as the lack of confident sensors, enforce the traffic engineers to further investigations.

The researches have direction to find simpler methods. They include directly measured traffic parameters to be used in determining the allowed number of vehicles to pass (through traffic signals), without including the prediction in real time. The artificial intelligence can offer a very different way to solve the above mentioned problems. The artificial intelligent system reaches the same results as man does when performing cognitive tasks. These systems are characterized with the ability to accumulate and use knowledge, set the problem, learn, conclude, solve the problem, process and exchange knowledge.

The machine learning is a field in artificial intelligence. The reinforcement learning is a technique within the machine learning. It is successfully applied in solving problems, e.g. operating the elevators and robot soccer games. It is also applied in modeling the supply chain, dynamic allocation of resources, predicting time series.

In this research the reinforcement learning will be applied in development of adaptive signal control strategy at an isolated intersection.

The paper is divided into two parts: the theory of the reinforcement learning is presented in the first part, while the methodology of the application of the Q learning in the development of adaptive signal control strategy at signalized intersection is presented in the second part.

The analysis of the results of the simulation with the application of this approach shows significant improvements in comparison with the traditional fully-actuated control.

2. ARTIFICIAL INTELLIGENCE

2.1. Reinforcement learning theory

The ability to learn is one of the characteristics that define the intelligence. That is why the machine learning is center point of the artificial intelligence.

The reinforcement learning is a subfield of the machine learning, learning what to do – how to map the state into actions and how and in which way to maximize the numerical reward. That is a learning driven by interaction with an environment and directed to the goal. It is a technique that does not need monitoring and it is different from the methods of the supervised learning (e.g. neural networks).

The learner or decisions maker is called agent, and everything it interact is called environment. The agent has a set of sensors to observe the state of the environment, and to be able to perform a set of actions to change the state of the environment. The most important characteristics of the agent are: trial and error search and delayed reward.

The learner or an autonomous agent that senses or acts in its environment can learn with trial to choose the optimal action or actions which lead to the highest reward.

For more accurate presentation of the interaction we can assume that the agent and the environment communicate in each sequence of discrete time steps: t=0,1,2,… In each time step, t, the agent receives some representation of the state of the environment, $s_t \in S$, where S is the set of possible states. In accordance with that, action $a_t \in A(s_t)$ is chosen, where $A(s_t)$ is a set of actions which are available in the state s_t. One step later, in part as a consequence of its action, the agent gets numerical reward, $r_{t+1} \in R$ and finds itself in a new state, S_{t+1}. The teacher can provide reward or penalty in order to induce the desirability of the final state. Figure 1 shows the agent-environment interaction.

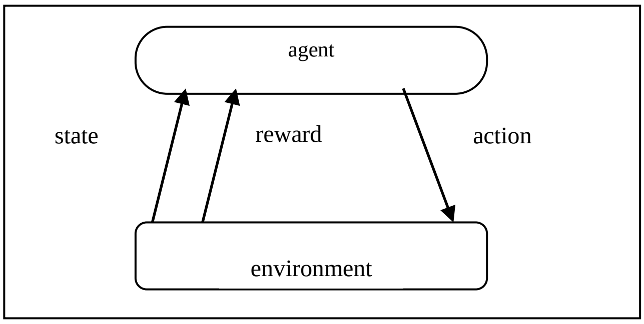

Fig. 1. Agent-environment interaction

The transit from one to another state is shown as:

$$S_0 \xrightarrow[r_0]{a_0} S_1 \xrightarrow[r_1]{a_1} S_2 \xrightarrow[r_2]{a_2}$$

Where
- S_i is the state in the time step i,
- a_i is the possible action available in each state in the time step i,
- r_i is the reward which the agent receives in the time step i for taking action a_i.

2.1.1 Reinforcement learning elements

The reinforcement learning has four basic elements, such as:

Policy – In each time step, the agent maps the presentation of the state with the probability of selection of each possible action. This mapping is called **agent policy** π_t, where $\pi_t(s,a)$ is the probability that $a_t=a$ if $s_t=s$. The reinforcement methods determine how the agent changes its policy as a result of its experience. The agent's goal, generally speaking, is to maximize the total reward it receives in long term.

Reward function – In the case of the reinforcement learning, the agent's goal is formalized in a sense of special signal, called **reward,** which is transmitted from the environment towards the agent. The reward is simply a number whose value differs from step to step. Informally, the agent's goal is to maximize the reward it receives. This means maximizing the reward that it receives in the moment as well as maximizing the cumulative reward that it receives in a long term.

Value function – Almost all reinforcement learning algorithms are based on the assessment of value functions – state functions (or state-action pairs) which assess how good is for the agent to be in a given state (or how good is to conduct some action in a given state). The term 'how good' is defined here in a sense of future rewards that can be expected, or, more precise, in a sense of expected earning (compensation). Of course, the rewards that the agent expect to receive in the future depend on the action it will conduct. Accordingly, the value functions are defined in accordance with some policies.

Having in mind that the agent learns directly from the interaction with the real environment, it does not need the **model of environment.** The agent can be taught previously by using a simulator. The use of a simulator helps the agent to avoid the unacceptable and degrading actions in the field. In a case of real time traffic control, this type of action can have an counter effect on the system in particular.

3. ADAPTIVE SIGNAL CONTROL STRATEGY DESIGN AT SIGNALIZED INTERSECTION

With the aim of gathering research data for this control strategy, simple network is created in a traffic micro simulator PTV Vision VISSIM COM.

As an input data it is used:

- Data for the traffic network – created graphically and
- Data for the traffic demand – entered through the graphical interface of the simulator as a profile of traffic demand in all zones – origins of traffic.

The data for measuring the efficiency of the control strategy are gathered directly with the assistance of the functions inside the VISSIM and saved in variables inside the program. The functions are called in each time step. The time step in use is 1 second. The program for communication with the simulator, gathering of the data and the reinforcement learning algorithm is done in VBA.

The design process is done in the following steps:

- Creation of traffic network with graphical implementation of the intersection, creation of links, connectors, traffic signals, setting detectors,
- Creation of traffic demand,
- Gathering and analysis of the results.

A four leg intersection is shown on Figure 2.

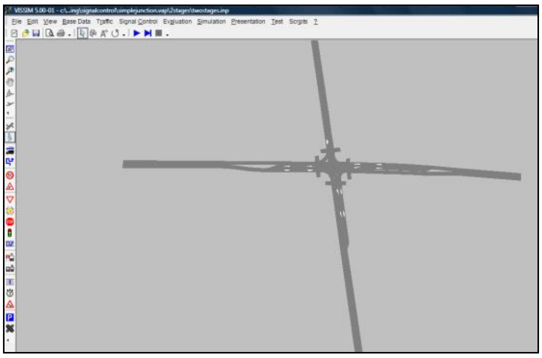

Fig. 2. Intersection design.

There are four zones for generation of traffic demand, one for each approach.

The user defines the name, the position and the dimensions of the detector with the help of the graphical interface of the simulator. Each detector is linked with appropriate traffic signal. After defining the detector, the data can be gathered in ASCII documents.

Powerful tool from the Intelligent transport systems that uses the technique of the artificial intelligence, so-called **reinforcement learning** – the Q-learning approach has been chosen to deal with the stochastic nature of the traffic and to develop adaptive signal control strategy at an intersection.

3.2 Q – learning elements

The Q – learning agent has three elements: **states (*s*), actions (*a*), rewards (*r*)**. In the developing strategy for the intersection shown in Fig. 2 we set:

States: phase, if it is a call for green extension; in this case the *gap* for the stages with green time and *occupancy* for the stages with red time are being examined.

Actions are defined as:
0 ...take action to change the phase
1...take action to continue the green phase.

The choice of action for this research is made according to ε - greedy policy where the best action is used with the possibility 1- ε and the research action is chosen by random choice with probability ε.

Reward: defined as a total number of vehicles at the intersection.

The algorithm of Q-learning (control algorithm of the temporary difference (off-policy TD Control)) is shown in its procedural form in the following way:

Initialize *Q(s,a)* arbitrarily
Repeat (for each episode):
  Initialize *s*
  Repeat (for each step of episode):
    Choose *a* from *s* using policy derived from *Q* (e.g., - greedy)
    Take action *a*, observe *r*, *s'*

$$Q(s,a) \longleftarrow Q(s,a) + \alpha[r + \gamma \max_{\alpha'} Q(s',a') - Q(s,a)]$$

S←s';

  until *s* is terminal.

In the algorithm above *Q (s,a)* is the function of the action value, α is the learning percentage, *r* is the reward, γ is the discount rate, $Q(s',a')$ is the value function for the new action a' and the new state s'. S is the set of possible states.

4. SIMULATION RESULTS

In order to evaluate the feasibility of the proposed adaptive strategy, the results obtained by the agent are compared with the results obtained by traditional actuated control.

The testing is performed after sufficient number of iterations with different values of the conditions and after converging the Q-values. The number of needed iterations for convergence depends on the size of the state – action space. The simulation is run during the peak hour. The efficiency of the strategy through delays is measured. In T.1 only a part of the results are shown – the comparison with the actuated control and Q learning.

T. 1: Comparative analysis of the performances for actuated control and Q learning strategy

Run	**Type of control**	**Delay (sec)**
278	Actuated	0,232619598507881
279	Actuated	0,175925269722939
280	Actuated	0,192265138030052
281	Actuated	0,1535364985466
282	Actuated	0,159261509776115
283	Actuated	0,249249652028084
284	Actuated	0,15443129837513
285	Actuated	0,163399875164032
278	Q - learning	0,187686145305634
279	Q - learning	0,152055725455284
280	Q - learning	0,182069316506386
281	Q - learning	0,161372631788254
282	Q - learning	0,17327706515789
283	Q - learning	0,191062957048416
284	Q - learning	0,1720094711
285	Q - learning	0,1137311101

According to the modest simulation results obtained, one can expect some promising and encouraging conclusions. Namely, some reductions in delays in a favor of Q-learning strategy are perceived. However, this hypothesis cannot be scientifically proven at this time because of the small number of iterations and the simplicity of the isolated intersection.

5. CONCLUSION

Although the number of applications of traffic control strategies is large, though this field is still a challenge for the researchers mainly because of the stochastic nature of the traffic flows.

In this paper there is an attempt to promote the application of adaptive control strategy at an intersection by using the technique of artificial intelligence known as reinforcement learning. Q-learning algorithm is used. The research is implemented by using VISSIM micro simulator, by direct programming of the functions in the simulator.

Encouraging results of the adaptive traffic signal control strategy to an isolated intersection are presented. The results of the simulation indicate that the new approach is more efficient compared to the traditional actuated control.

Further valuations of the strategy under different conditions and different settings of the elements of the Q learning agent will be additionally developed.

1. REFERENCES

[1] Sutton, R.S., Barto, A.G., Reinforcement Learning - An Introduction. MIT Press, Cambridge, Massachusetts, 1998.

[2] Russell, S., Norvi, P., Artificial Intelligence: A Modern Approach, Prentice Hall, 2009

[3] Abdulhai B., Kattan L., “Reinforcement learning: Introduction to theory and potential for transport applications, Can. J. Civ Eng. 30, pp 981-991, 2003.

[4] Abdulhai, B., Pringle, R., Karakoulas, G.J., Reinforcement Learning for True Adaptive Traffic Signal Control, ASCE Journal of Transportation Engineering, Volume 129, Number 3, pp278-285, 2003

[5] Shi, Z., Principles of Machine Learning, International Academic Publishers, Beijing, 1992.

[6] Veljanovska, K., Bombol, K., Choosing a new tool for adaptive control strategy design, PhD Research day RESEARCH PROCEEDNIGS, May 9, 2007.

[7] Spall, J. C., Chin, D., Traffic -Responsive Signal Timing for System-Wide Traffic Control, Transportation Research Part C, Vol. 5, No. 3/4, Elsevier Science Ltd, Great Britain, pp. 153-163,1997.

[8] Wiering, M., et all, Intelligent Traffic Light Control, Institute of information and computing sciences, Utrecht University, Technical report UU-CS-2004-029.

[9] Roozemond, D., Veer, P. V. D., Usability of Intelligent Agent Systems in Urban Traffic Control, Delft University of Technology, Delft, 1999.

[10] Meystel, A & Messina E., The Challenge Of Intelligent Systems, Proceedings of the 15th, IEEE International Symposium on Intelligent Control (ISIC 2000), Rio, Patras, GREECE 17-19 July, 2000.

[11] Shoufeng, L., et all, Q – Learning for Adaptive Traffic Signal Control Based on Delay Minimizations Strategy, IEEE International Conference on Networking Sensing and Control, pp 687 – 691, 2008.

[12] Koltovska, D., Bombol, K., Analysis of the Potentials of the Artificial Intelligence Techniques for Adaptive Signal Control, ISEP 2009, R 2, Ljubljana, 2009

[13] VISSIM 5. 00 User manual, PTV Planung Transport Verkehr AG, Germany, 2007

[14] Roca, V., VISSIM COM – User Manual for the VISSIM COM Interface – VISSIM 5.0, PTV Planung Transport Verkehr AG, Karlsruhe, 2007.

ADAPTIVNO UPRAVLJANJE SVJETLOSNIM SIGNALIMA POMOĆU TEHNIKA UMJETNE INTELIGENCIJE

Daniela Koltovska Nečoska
Kristi Bombol
Sveičilište Sv. Kliment Ohridski, Tehnički Fakultet
Odjel za Promet
POB 99, 7 000 Bitola, Republika Makedonija

Sažetak

Veliki je broj inteligentnih sustava koji se mogu prilagoditi promjenama prometnog toka i biti sposobni za učinkovito upravljanje prometom i izmjenu signalnih pojmova.

Dugo se vremena smatralo da će sustavi u realnom vremenu omogućiti značajne prednosti. No međutim, postoje brojna ograničenja kao što je primjerice korištenje modela s visokom razinom preciznosti detalja. Konvencionalne metode optimalnog upravljanja pate od takozvanog prokletstva dimenzije. Poteškoće u optimalnom upravljanju svjetlosnim signalima i važnost problema doveli su do brojna nova istraživanja. Moguća se rješenja navedenih problema naziru u metodama umjetne inteligencije. Umjetni inteligentni sustav postiže iste rezultate kao i čovjek kod obavljanja kognitivnih zadataka. Ovi sustavi imaju sposobnost da akumuliraju i koriste znanje, postave problem, uče, zaključuju, riješavaju problem, kao i da razmjenuju znanje.

U istraživanju prikazanom u ovom radu prijedlaže se primjena adaptivnog upravljanja svjetlosnim signalima pomoću inteligentnog agenta kontrole koji je sposoban prilagoditi optimalnu politiku učenjem iz okruženja. Cilj koj ise treba postići je smanjenje vremenskih gubitaka u sustavu.

U radu će biti prezentirano postavljanje algoritma Q-učenja i prvi rezultati primjene u razvoju adaptivnog upravljanja svjetlosnim signalima. Pokazuje se da su dobiveni rezultati glede adaptivnog upravljanja u prijednosti u odnosu na automatsko upravljanje prometom.