METHODOLOGY FRAMEWORK FOR DEVELOPING ADAPTIVE TRAFFIC CONTROL STRATEGY – A NOVEL CONCEPT FOR TRAFFIC ENGINEERS

Assistant Daniela Koltovska, MSc
Prof. Kristi Bombol, PhD

Sv. Kliment Ohridski University
Faculty of Technical Sciences
Department for Traffic and Transport
POB 99, 7 000 Bitola, Republic of Macedonia
E-mail: kristi.bombol@tfb.uklo.edu.mk
daniela.koltovska@tfb.uklo.edu.mk

Abstract

The intelligent agent technology, sharing the characteristics of applicability in real time and adaptation, capability of self-analyzing by errors and success, learning and improvement (in the course of time) interacting with the environment, quick learning from a large amount of data, represents the new approach employed in the development and design of new adaptive control strategies. These are strategies that incorporate a higher level of intelligence and are capable of self-learning and experience-based decision making.

However, the process of developing such strategies, that are undoubtedly our future, is by no means easy, particularly as regards traffic engineers – researchers. Nowadays they are facing a double challenge: application of new unknown, narrowly specialized methods in the area of computer science, as well as accessibility to professional literature, predominantly meant for individuals possessing pre-conception in the area of artificial intelligence and machine learning.

Hence is our motive to write this paper – to present the methodology for designing and developing an adaptive control strategy for an isolated intersection, by applying the intelligent agent and the reinforcement learning method, refracted through the traffic engineer's prism. The designed methodology comprises of three steps: development of a model, design and development of an intelligent agent, strategy testing and evaluation. Each step is explained in an easily understandable manner from the point of view of a traffic professional.

We assume that this clearly established methodology is going to be of invaluable assistance to all traffic engineering researchers dealing with the issue of artificial intelligence and machine learning for the first time.

Key words: adaptive control, strategy, intelligent agent, reinforcement learning, Q learning, intersection

INTRODUCTION

The emergence of the intelligent agent concept, a significant move in the overall information science has been made. Nowadays this concept is applied in traffic when developing adaptive control strategies. The idea behind is the autonomous entities known as agents to start learning to behave in an optimal way by direct interaction with the system. By applying machine learning (ML) algorithms that are based on rewards or penalties depending on the results obtained in the actions selected by the agent, the optimal policy trying to optimize the traffic flow can be calculated.

In order to develop an adaptive control strategy, first it is necessary to set an appropriate clear-cut methodology. This paper is aiming at presenting the methodological approach to designing and developing the adaptive control strategy of an intersection, applying the reinforcement learning (RL) method. The steps to development of the control strategy, such as, identifying the type of intersection, the development of a model and implementation of Q-learning algorithm, data and tools used in the development of an intelligent agent (IA), are precisely described.

The control strategy presented in this paper is performed by an agent. In order to embed the learning feature in the agent, the RL method is applied, as well as the Q- learning algorithm.

METHODOLOGY DESCRIPTION

The process of developing adaptive control strategy at an isolated intersection, is composed of three steps (Figure 1):

Step 1: **Development of model**

Step 2: **Design and development of the Intelligent Agent (IA)**

Step 3: **Strategy Testing and Evaluation**

Further in the text each step is going to be described in clear manner.

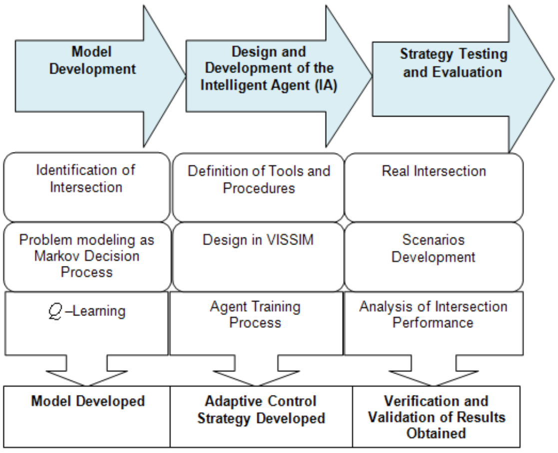

Figure 1: Methodology of adaptive control strategy development process
Source: developed by the author

MODELING THE TRAFFIC SIGNAL CONTROL APPLYING THE RL METHOD

Identification of intersection type

At the very beginning it is indispensable to determine the type of intersection from the geometry characteristics point of view. For the research purposes, a four-leg intersection with a left-turn lane has been selected. For the selected type, control logic has been designed, the cycle length has been defined, as well as the number of phases, the minimal and maximal green for each signal groups within the phase.

Mathematical interpretation of the RL process

When Markov decision process exists, the process of RL can be applied. In this context, to have the RL agent learn the control policy (to take decisions for changing the traffic signal states), it is necessary to determine the *set of states*.

Defining the set of states S (from the research example)

The selection of the variables to describe the traffic process greatly varies. Within the research, *phase, gap, occupancy* are applied.

$S = \{(\phi, g, Occ);\ \ \phi \in \{1,2\},\ \ g \in \{YES, NO\},\ Occ \in \{0,1\}\}$, where ϕ is the signal phase within a signal cycle of C = 90 seconds); when ϕ = 1 it is a green phase, when ϕ = 2, the phase is red. Green time t_g, within a single signal cycle C of 90 seconds falls within the interval of 24 to 78 seconds, i.e. $t_g \in [24,78]$. Red time t_r within a single signal cycle C of 90 seconds falls within the interval of 12 to 66 seconds, i.e. $t_r \in [12,66]$. g is a binary variable receiving the values $\{YES, NO\}$, where the value of NO denotes that there are no vehicles (signal received from the inductive loop); YES represents the opposite. Occ is a binary variable, where the value of 0, denotes that there are no present vehicles from the conflict flow (red light), and the value of 1, denotes the opposite.

Defining the set of actions A

Based on the information related to the detected state, the control agent takes up action. For each state, the agent can only take up two actions: action value of 1, which means the state remains the same (green time extension), or action value of 0, which means change of the signal state.

Defining the set of rewards R

The rewarding function is the second key element for the agent. The reward is a function that depends on the system's state and the action taken. The reward takes values from the set of natural numbers, i.e. it is defined as mapping $R: A \times S \rightarrow \subseteq$.

The rewarding function goal is maximization of the total throughput.

As the research problem is mathematically explained as Markov decision process (MDP), the *Q*-learning is applied.

Q-learning

The recursive definition of Q-function provides iterative algorithm for deducing the Q-function. The procedure to achieve this is as follows:

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.

Each action, derived by the agent, influences the environment; upon the completion of the action, the environment is in a new state. For each action taken, the agent is rewarded and the reward defines the extent to which the action was good or bad. The rewarding helps the agent to learn what to do and to act in a *more intelligent manner*.

DESIGN AND DEVELOPMENT OF THE INTELLIGENT AGENT (IA)

Defining the tools and procedures

In the design of the intelligent agent, the following tools have been used:

- *VISSIM 5.4-0.3 – (**V**erkehr **I**n **S**tädten - **SIM**ulationsmodell; "Traffic in cities - simulation model")*
- *VISSIM **COM** (**C**OMPONENT **O**BJECT **M**ODEL)*
- *Microsoft SQL (MSSQL)*
- ***S**IDRA INTERSECTION 5.1 – (Signalised & unsignalised **I**ntersections **D**esign and **R**esearch **A**id).*

Design by VISSIM simulator

To learn the control strategy, the RL agent requires a simulated traffic system environment. The simulation platform that is being used is VISSIM. The traffic demand has been created via the simulator's graphical interface. The number of vehicles is entered for every link at intervals of 15 minutes per peak hour (known/unknown demand). Vehicle arrivals are described by the Poisson distribution.

As the strategy applies for the peak hour, the simulation period is 1 hour (3600 seconds). To express the stochastic variations of traffic flows as realistically as possible, the parameter used to initialize a random number generator is applied (*Random Seed*).

From the user side, the number, the position and the detector dimension are defined by applying the simulator's graphic interface. Each detector is connected with a corresponding signal and a corresponding phase.

The program for communication among VISSIM simulator, the database and the RL algorithm is developed using the C Sharp (C#) program language (Figure 2).

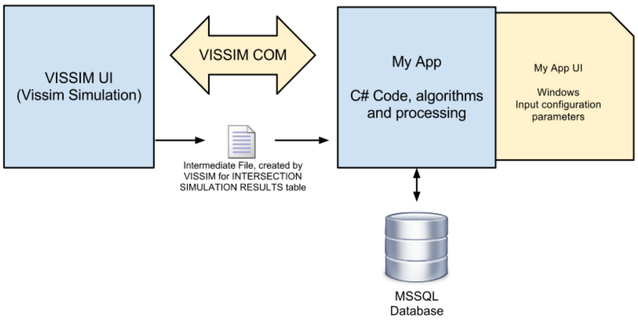

Figure 2: The process of communicating and interaction among the main elements
Source: developed by the author

The agent is being trained in simulation conditions. However, after being applied in the field, the agent can continue to learn, starting with the last Q – values obtained in the training process.

After sufficient number of iterations and convergence of Q – values, the training phase is completed. The next phase is testing and evaluation the adaptive control strategy.

STRATEGY TESTING AND EVALUATION

The strategy testing is performed on a real four leg intersection located within the central area of Bitola. Picture 2 depicts the intersection and the communication with the RL intelligent agent.

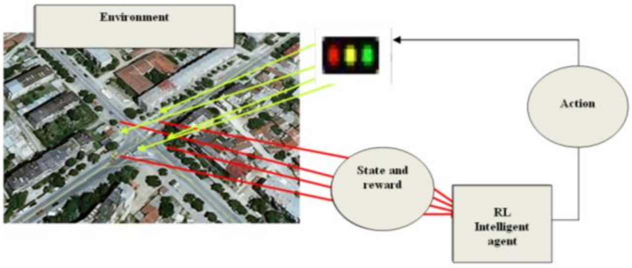

Figure 3: Description of intersection and communicating with the RL agent
Source: developed by the author

Delay, throughput and *number of stops* are analyzed as strategy efficiency meaures.

The results obtained from the learning IA are compared to the ones obtained through simulations in cases of fixed time and actuated control. The fixed time control is selected as a base case and all the other results are estimated in relation to it.

The testing is performed after a sufficient number of iterations with various values regarding states and after the convergence of Q–values. When testing, the selected action is the one with maximum Q value and the one that will provide optimum control action in all of the agent states.

Depending on the traffic flow conditions, and whether the traffic demand is known or unknown to the agent, the testing is performed in two phases. During the first phase, the testing is performed for uncongested traffic conditions with known and unknown demand. During the second phase, the testing is performed for congested traffic conditions with known and unknown demand.

The testing has shown that the adaptive strategy is superior when compared to the base case (fixed time control) for all phases and scenarios. In relation to actuated control, improvements are perceived for congested traffic conditions and balanced high demand flows on all intersection approaches.

CONCLUSIONS

This paper is attempting to present a clear and unambiguous methodological frame for developing adaptive control strategy for an isolated intersection. The aim was to produce an easily understandable guide to the sequence of activities – where to start and which a researcher is to follow, when entering the area of artificial intelligence and lacking sufficient computation background.

The designed methodology comprises three steps: development of model, design and development of the intelligent agent, strategy testing and evaluation. Each step is described in a comprehensible manner from the traffic expert point of view.

We assume that this clearly established methodology is going to be of invaluable assistance to all traffic engineering researchers dealing with the issue of artificial intelligence and machine learning for the first time.

REFERENCES

1. Abdulhai, B. and Pringle, R., Autonomous Multiagent Reinforcement Learning – 5 GC Urban Traffic Control?, A paper presented at the 2003 Annual *Transportation Research Board Meeting*, Washington, D.C, 2003.
2. Abdulhai, B., and Kattan, L., Reinforcement learning: Introduction to theory and potential for transport applications, *Canadian Journal of Civil Engineering*, Volume 30, 129(3), 2003, pp 981–991.
3. Abdulhai, B., Pringle, R., and Karakoulas, G., Reinforcement Learning for True Adaptive Signal Control, *Journal of Transportation Engineering*, Volume 129, Number 3, 2003, pp. 278-285.
4. Abdulhai, B., Pringle, R., and Karakoulas, G., Reinforcement Learning for ITS: Introduction and a Case Study on Adaptive Traffic Signal Control, *Transportation Research Board*, Washington D.C., 2001.
5. Abdulhai, B., What is After Vehicles Are Connected: How Do We Control Traffic and Congestion? Toronto ITS Centre, University of Toronto, Available at http://www.ece.ubc.ca, accessed on 13.10.2012.
6. Russel, S., Norvig, P., Artificial Intelligence, A modern Approach, Third Edition, Prentice Hall, 2010.
7. Sutton, R.S., Barto, A.G., Reinforcement Learning - An Introduction. MIT Press, Cambridge, Massachusetts, 1998.
8. Bombol K., Koltovska D., Veljanovska K., "Application of Reinforcement Learning as a Tool of Adaptive Traffic Signal Control on Isolated Intersections", *International Conference on Traffic and Transportation Engineering* (ICTTE 2011), ICTTE 2011, pp 75-79 Dubai