DOI 10.20544/HORIZONS.B.03.1.16.P17
UDC 656.1.05:004.8

INTELLIGENT TRAFFIC CONTROL IN URBAN AREAS[1]

Edouard Ivanjko
Department of Intelligent Transportation Systems,
Faculty of Transport and Traffic Sciences, University of Zagreb, Vukelićeva 4
Zagreb, Republic of Croatia
edouard.ivanjko@fpz.hr

Daniela Koltovska Nechoska
Faculty of Technical Sciences
St. Kliment Ohridski University, Bitola, Blvd 1st May, nn
Bitola, Republic of Macedonia
daniela.koltovska@tfb.uklo.edu.mk

Abstract

Control systems responding to the current traffic situation by adapting its parameters enable significant benefits. However, numerous limitations exist such as the need for accurate traffic models, the uncertainty in predicting future traffic flows, the difficulty in arrival time estimation, and the lack of a self-adjusting mechanism. Difficulties in optimising the signal control strategy and the importance of finding a solution to this problem resulted in different approaches. Methods from artificial intelligence have emerged as one possible solution. Such methods have the ability to accumulate and use knowledge, set a problem, learn, process, conclude, solve the problem and exchange knowledge. In this paper, the application of the Q-learning algorithm to control an independent intersection and an on-ramp on an urban motorway is described. Implemented approach enables an on-line adaptation of the control parameters to the current traffic situation reducing the delay in traffic.

Keywords - urban traffic control; artificial intelligence; Q learning; ramp metering; independent intersections

---

[1] Original scientific paper

INTRODUCTION

Urban areas are today prone to significant reoccurring traffic congestions causing problems related to delays in goods delivery, prolonged travel time, increased pollution and reduced quality of life. To solve this problem many cities change their policy regarding transport and subsidize mode shift from cars to public transport and bicycles or the use of electric vehicles. Additionally, urban traffic control centres are built to collect traffic data and manage the complete urban traffic network as a whole. To achieve an improvement in travel quality appropriate traffic control algorithms are needed in such traffic control centres.

In the recent decade, algorithms from the domain of artificial intelligence are being used [1]. They have the ability to accumulate and use knowledge, set a problem, learn, process, conclude, solve the problem and exchange knowledge with other systems. Therefore, management of a large urban traffic network becomes feasible and additional criteria like pollution levels can be added to the classic traffic related criteria (travel time, delay, queue length, etc.). In this paper, the potential of one such algorithm from the domain of reinforcement learning for application in urban traffic control is described. The Q-learning algorithm is chosen because of its suitability to control traffic processes, which can be described with a set of states and can be influenced by a limited number of control actions. Additionally, the Q-learning based traffic controller can be implemented as an intelligent agent enabling so an easier information interchange between different traffic controllers. Two application cases in the urban environment are used to demonstrate the effectiveness of the Q-learning algorithm. First case is related to control of an independent intersection and the second one to control of a local on-ramp on an urban motorway.

TRAFFIC CONTROL PROBLEMS IN URBAN ENVIRONMENTS

Urban environments are prone to heavy daily congestion periods known as rush hours, there is a lack of space for transport infrastructure build-up (especially in dense populated areas), lack of parking spaces creates additional traffic demand, contain several modes of transport, contain connections between the local road network and urban motorways, serve local and transit traffic, etc. The challenge is how to optimally control one part of the road network and not to cause problems in surrounding parts of the road network. To overcome this challenge, the problem of local control has to be solved first.

When traffic control problems in urban environments are examined, one can distinguish two parts of the urban road network. First part is the road network in the central urban area and the second part is the surrounding road network. Traffic in both mentioned parts of the road network can be modelled using wave theory. According to the wave theory, vehicles travel in groups creating longitudinal waves. When such a group of vehicles arrives, congestion can occur if vehicles of surrounding roads are merged into the corresponding traffic direction without any restrictions. Better solution is to let the wave pass using the surrounding roads as temporary vehicle storage places. This can be accomplished by using appropriate control approaches.

To control the traffic in the central urban network traffic lights are mostly used on intersections. For optimal control of intersections, appropriate signal plans for traffic lights have to be generated. Best-suited approach for this is an adaptive approach with the ability to react on the changing traffic demand [1]. To control the surrounding road network additional to traffic lights variable message signs are used. Such signs have the ability to inform the drivers about a traffic congestion ahead, about dangerous road surface conditions, to impose different speed limits, etc. In this paper only the possibility to change the signal plans of traffic lights according to the current traffic situation is examined. According to this in the central part signal plans for intersections can be changed and in the surrounding part, the signal plans or ramp metering rate for the traffic light on on-ramps can be changed.

THE Q-LEARNING ALGORITHM

As mentioned above, effective control of traffic flows with variable traffic demand needs approaches, which can take into account the changing traffic situation. Variable traffic demand is a result of random choices of traffic users and computed choices of a traffic controller. The traffic controller can in each time step *n*change the traffic state*s* by applying a control input action *a* to the current traffic state [2]:

$$s_{n+1} \sim p\left(s_{n+1} \mid s_n, a_n\right), \qquad (1)$$

where *p* is a probability distribution function over the state action space. All processes, which can be described with (1) are Markov decision processes (MDP) and the belonging probability distribution function represents the Markov model of the whole system. In this paper, the whole system contains the traffic controller and the underlying traffic process to be controlled.

MDPs can be described by a 5-tuple$\left(S, A, P, R, \gamma\right)$ where *S* is a finite set of states ($s \in S$), *A* is finite set of actions ($a \in A$), *P* presents the transition probability from a particular state s_n to a new state s_{n+1} if action a_n has been taken, R presents the reward received from the state transition, and γ is the discount factor ($\gamma \in [0,1]$). The discount factor γ represents the difference in importance between future and present rewards. The reward function depends on the chosen action or of the so-called policy function $\pi(s)$ applied on a particular state ($r(s,\pi(s)) \in R$). Therefore, the problem of controlling a MDP can be defined as a problem to find the appropriate policy function that an intelligent agent (traffic controller in this case) will apply to choose the optimal action for the transition from state s_n to s_{n+1}.

One of the basic approaches to learn the needed policy function is reinforcement learning (RL). When RL is applied, the traffic controller is implemented as an intelligent agent and it enables the agent to work in a framework and gain new knowledge during operation. In this paper, the Q-learning algorithm is used to learn the policy function of the agent. Since in this paper the agent is a traffic controller, the Q-learning algorithm will learn the optimal control law for the underlying traffic process. To learn the needed control law the following learning rule is applied:

$$\hat{Q}_n(s,a) \longleftarrow (1-\alpha_n)\hat{Q}_{n-1}(s,a) + \alpha_n[r + \gamma \max_{\alpha'} \hat{Q}_{n-1}(s',a')], \qquad (2)$$

where the learning rate α_n is defined as:

$$\alpha_n = \frac{1}{1 + visits_n(s,a)}\ , \qquad (3)$$

and where $\hat{Q}_n(s,a)$ is the expected value of the previous defined value for a deterministic function case for an action *a* and state *s*, $\hat{Q}_{n-1}(s',a')$ is the expected value of the previous defined value for the new action *a'* in the next state *s'*, α_n is the learning rate, *(s,a)* presents the updated state and action during *n* time steps, and $visits_n$ *(s,a)* is the total number of visits for a state-action pair until the n^{th} time step.

Appropriate states, actions and the reward function have to be defined according to the specifics of the underlying process to be controlled. Since the underlying process is a road traffic process, appropriate states are related to traffic parameters that can clearly describe the current traffic situation, and are different for the case of an independent intersection [3] and for the case of an urban motorway [4, 5]. Set of actions incorporates changes of signal plans for traffic lights and the reward function is related to traffic parameters describing the throughput of the controlled traffic network. In continuation more details about the implementation of the Q-learning algorithm for control of the two mentioned cases are given.

CONTROL OF INDEPENDENT INTERSECTIONS

Independent intersections present the first case of application of the Q-learning algorithm for traffic control.The set of states $S$was defined as [3]:

$$S = \{(\phi, g, Occ); \phi \in \{1,2\}, g \in \{YES, NO\}, Occ \in \{0,1\}\}, \quad (4)$$

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

Based on the information related to the detected state, the control agent choses an appropriate action. For each state, the agent can chose between 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.The rewarding function is the second key element for the agent. In this case the goal of the rewarding function is maximization of the total throughput. For that purpose, the following set of rewards was defined:

1.) Reward Function –total throughput;
2.) Immediate reward –number of vehicles passing at green light in the previous time interval(the length of this interval is 90 seconds);
3.) Discounted reward – total number of vehicles in the peak hour.

The action is taken at a shorter interval for a given time step. The vehicles are counted per one 90 second signal cycle. An action is taken per second – over the green time duration, in which case the step takes 3 seconds.

LOCAL RAMP METERING

Local ramp metering presents the second case of application of the Q-learning algorithm for traffic control. The basic concept of local ramp metering is given in Fig. 1. Term local denotes in this case that only one on-ramp is controlled. The inflow to the urban motorway is controlled by the means of a special traffic light that contains only the red and green phase. Duration of the green light is fixed and enables only one or two vehicles to merge the mainstream traffic flow. Duration of the red light can be changed defining so the amount of vehicles that can merge with the mainstream traffic flow. Vehicles are temporally stored on the on-ramp so the queue length has to be monitored to prevent a congestion spillback into the urban arterial road network.

Many approaches exist to ramp metering and today the emphasis is on implementing intelligent control approaches [1]. Q-learning is also one of intelligent based approaches that can be implemented for ramp metering. For the implementation appropriate states, actions and reward function have to be defined. In order to enable good level of service (LoS) on the urban motorway in cases of significant traffic demand change the following states have been chosen [5]:

$$S = \{(\phi, \rho, q); \quad \phi \in \{1,2,3\}, \quad \rho \in \{0,1,2\}, \quad q \in \{0,1,2\}\}, \qquad (5)$$

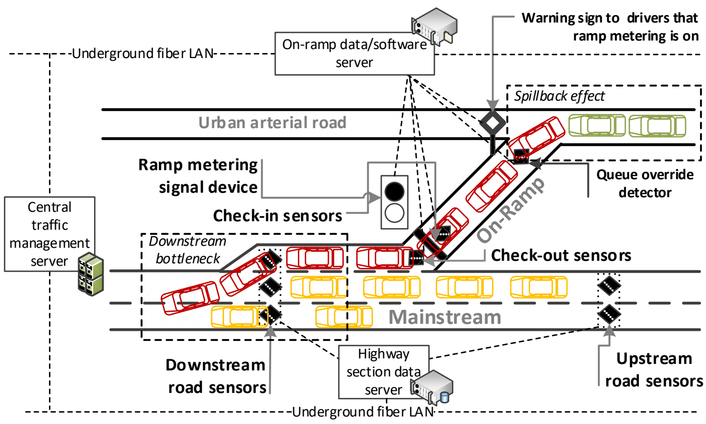

Fig.1. Traffic control problems on urban motorways and infrastructure for local ramp metering [4]

where ϕ is the signal phase, ρ is density of the downstream mainstream traffic flow, and q is the on ramp queue length. Detailed description of the mentioned states and representative values can be found in Table 1.

Table 1. Description of all states for local ramp metering [5]

States	Values	Description
Phases	1	Represents the "all green" phase with fixed duration of 3 s (one vehicle per green strategy)

	2	Represents the "all red" phase calculated by the ramp metering algorithm (extension of current phase duration)
	3	Represents the "ramp metering off" phase which is activated in the case of low mainstream density
AverageDensity Class	0	Downstream density is between 0 [veh/km] and 100 [veh/km]
	1	Downstream density is between 101[veh/km] and 350 [veh/km]
	2	Downstream density is larger than 351 [veh/km]
AverageQueue Class	0	On-ramp queue length is between 0 and 4 vehicles
	1	On-ramp queue length is between 5 and 7 vehicles
	2	On-ramp queue length is larger than 8 vehicles

The agent makes a decision, which action to apply on the traffic process, every 3 seconds. To cope with significant changes in traffic demand two different sets of actions were defined. For each state, the agent can only decide between two actions in the case of local ramp metering in the first set of actions. First action is denoted by the value 1 and it indicates a traffic situation for which it is necessary to stay in the current signal phase. Second action is denoted by the value 2 and it indicates a traffic situation for which it is necessary to change the current traffic signal phase. The set of actions consist of two values only ($A \in \{1,2\}$) modelling the traffic signal phase change. Since for ramp metering only the green and red traffic light phases are used, two actions are enough. In the second set of actions, a third action was added denoting the turn off phase of ramp metering. This action is triggered if a longer period of low traffic demand on the mainstream and the on-ramp exist.

The reward function is in this case related to the on-ramp queue length and mainstream density. So, it is ensured that a period with low traffic demand (no vehicles waiting on the on-ramp) and periods with high traffic demand(long queue on the on-ramp and high density in the mainstream) can be detected. Reward for a particular traffic solution is added when the on-ramp queue category has the value 0 or/and 1. Additional reward is added in when mainstream density reaches category 0.

SIMULATION RESULTS

To simulate the two cases the microscopic VISSIM simulator is used. The Q-learning algorithm is implemented as a separate application in scope of the framework given in Fig. 2. In this framework the left part presents the traffic model and the right part present the control part. Both parts are connected using the VISSIM COM interface to interchange traffic sensor data and control inputs for traffic light signal plan changes.

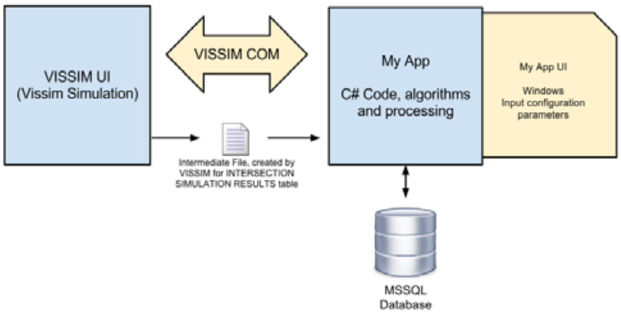

Fig.2. Implemented simulation framework [3]

Local intersection and Q-learning

The Q-learning based signal control testing is performed on a real four-leg intersection located within the central area of Bitola, using real traffic data. Figure 3 depicts the intersection and the communication with the RL intelligent agent.

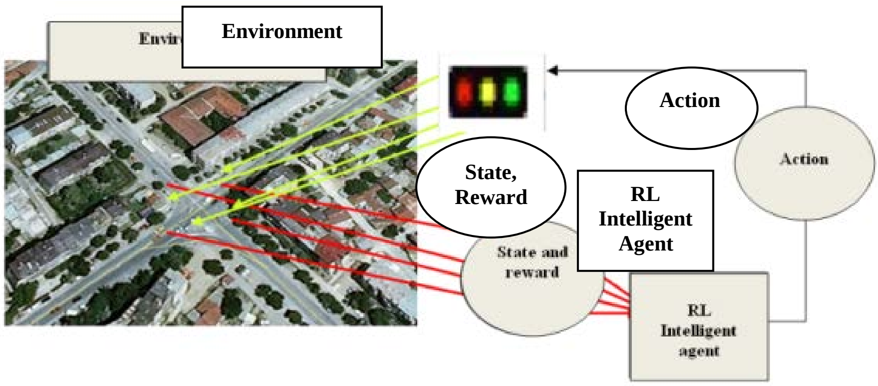

Fig.3. Description of intersection and communication with the RL agent [3]

Delay, throughput and number of stops are analyzed as efficiency measures for all types of intersection control. The results obtained from the learning intelligent agent are compared to the ones obtained through simulations in cases of fixed time and actuated intersection 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 three hundred of iterations with various values regarding states and after the convergence of Q–values. During testing, the selected action is the one with the maximal Q value. Such chosen action corresponds to the current 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. Figure4 shows the comparison of percentage of efficiency measure improvements for all phases of testing in case of applying fixed-time to that of Q-learning based signal control.

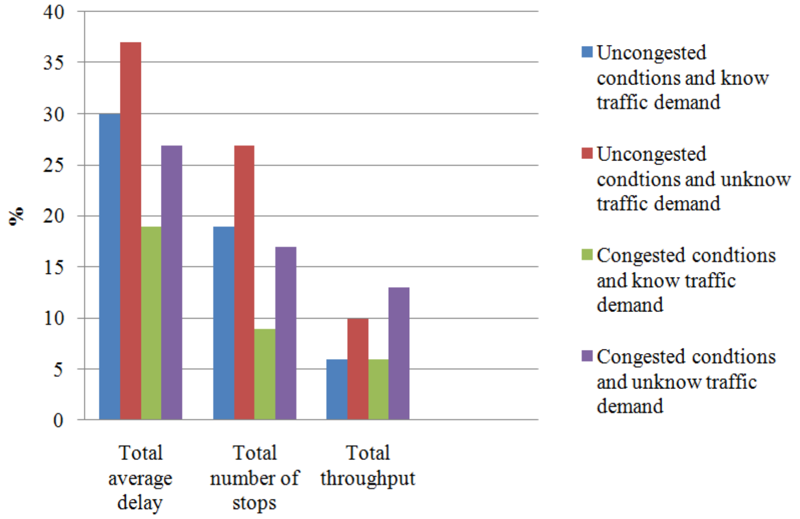

Fig.4. Comparison of percentage of efficiency measure improvements (fixed-time / Q-learning base signal control)

Overall, the following has been observed:

– The Q-learning based signal control gives best results with total average delay (improvement of 37%) and with a total number of stops (improvement of 27%), *in uncongested traffic conditions for unknown traffic demand*;

– The Q learning based signal control gives best results with the total throughput (13%) *in congested traffic conditions for unknown traffic demand.*

LOCAL RAMP METERING AND Q-LEARNING

To test the implemented local ramp metering algorithm a motorway traffic model with one on-ramp has been made. Duration of the simulation was set to 6 hours and during the first 2 hours there was low traffic demand, and during the next 4 hours there are 2 significant increases and decreases of traffic demand [5].Table 2 contains the simulation results for local ramp metering. It can be noticed that application of ramp metering can increase the LoS on the mainstream but causes delays on the on-ramp. Delays on the on-ramp reduce the LoS especially in cases with low traffic demand when ramp metering causes an unnecessary delay on the on-ramp. In the case where the second set of actions is used, one can notice further improvement of the LoS on the mainstream but also on the on-ramp compared to the results obtained with the first set of actions.

Table 2. Performance measures for local ramp metering

Performance Measures	**Control Method**		
	No ramp metering	**Q learning**	
		First set	**Second set**
Average Mainstream Travel Time (s)	61.9	58.2	57.4
Average On-ramp Travel Time (s)	97.4	182.9	100.4
Average downstream Speed (km/h)	82.1	87.3	88.5
Average on-ramp Speed (km/h)	30.0	16	29.1
Average Speed on the mainstream (km/h)	70.19	61.56	73.61
Total Delay (h)	60.01	88.07	50.76
Total Travel Time (h)	200.84	228.88	191.81

CONCLUSION

In the recent decade, algorithms from the domain of artificial intelligence are being used for traffic control [1]. They have the ability to accumulate and use knowledge, set a problem, learn, process, conclude, solve the problem and exchange knowledge with other systems.In this paper, the potential of one such algorithm from the domain of RL for application in urban traffic control is described. Two application cases in the urban environment have been used to demonstrate the effectiveness of the Q-learning algorithm. First case is related to control of an independent intersection and the second one to control of a local on-ramp on an urban motorway.

Simulation results from both test cases demonstrate that a Q-learning based traffic signal control can achieve evident improvements regarding chosen efficiency measures including unknown traffic demand in over-capacity congested traffic conditions. Future work on this topic will include augmentation of the Q-learning framework for networked traffic signal systems and integration with dynamic route guidance.

REFERENCES

[1] Mun Ng, K., Bin Ibne Reaz, M., Alauddin Mohd Ali, M., Guy Chang, T., "A brief survey on advances of control and intelligent systems methods for traffic-responsive control of urban networks", Technical Gazette, Vol. 20, No. 3, pp. 555-562, 2013

[2] Davarynejad, M., Hegyi, A., Vrancken, J., van den Berg, J., "Motorway Ramp-Metering Control with Queuing Consideration using Q-Learning", presented on the 14th International IEEE Conference on Intelligent Transportation Systems, Washington, DC, USA, October 5-7, pp. 1652-1658, 2011.

[3] Koltovska, D., Bombol, K.: "Intelligent Agent Based Traffic Signal Control on Isolated Intersections", TEM Journal, Vol. 3, No. 3. pp. 216-222, 2014.

[4] Ivanjko, E., Koltovska Nečoska, D., Gregurić, M., Vujić, M., Jurković, G., Mandžuka, S., "Ramp metering control based on the Q-learning algorithm", Cybernetics and Information Technologies, Special Issue on Control in Transportation Systems, Vol. 15 (2015), No. 5; pp. 88-97, DOI: 10.1515/cait-2015-0019

[5] Gregurić, M., Koltovska Nečoska, D., Ivanjko, E., "Application of Q-Learning to Ramp Metering in Cases of Significant Changes in Traffic

Demand", Autonomic Road Transport Support systems: Early Career Researcher Conference, University of Malta, La Valletta, Malta 27 - 28 May 2015.