Transportation Research Part C 54 (2015) 40–55

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A modified reinforcement learning algorithm for solving coordinated signalized networks

Cenk Ozan, Ozgur Baskan, Soner Haldenbilen, Halim Ceylan *

Department of Civil Engineering, Engineering Faculty, Pamukkale University, Denizli 20070, Turkey

ARTICLE INFO

Article history:
Received 11 November 2013
Received in revised form 8 March 2015
Accepted 8 March 2015
Available online 22 March 2015

Keywords:
Reinforcement learning
Coordinated signalized network
TRANSYT-7F
Signal timing optimization

ABSTRACT

This study proposes Reinforcement Learning (RL) based algorithm for finding optimum signal timings in Coordinated Signalized Networks (CSN) for fixed set of link flows. For this purpose, **MO**dified **RE**inforcement **L**earning algorithm with **TRANS**YT-7F (MORELTRANS) model is proposed by way of combining RL algorithm and TRANSYT-7F. The modified RL differs from other RL algorithms since it takes advantage of the best solution obtained from the previous learning episode by generating a sub-environment at each learning episode as the same size of original environment. On the other hand, TRANSYT-7F traffic model is used in order to determine network performance index, namely disutility index. Numerical application is conducted on medium sized coordinated signalized road network. Results indicated that the MORELTRANS produced slightly better results than the GA in signal timing optimization in terms of objective function value while it outperformed than the HC. In order to show the capability of the proposed model for heavy demand condition, two cases in which link flows are increased by 20% and 50% with respect to the base case are considered. It is found that the MORELTRANS is able to reach good solutions for signal timing optimization even if demand became increased.

1. Introduction

The optimization of traffic signal timings has been at the heart of urban traffic control for many years. It is well known that traffic signal control which encloses delay, queuing, pollution and fuel consumption is a multi-objective optimization. For a signal controlled road network, using the optimization techniques in determining signal timings has been discussed greatly for decades. Due to complexity of signal timing optimization problem, new methods and approaches are needed to improve efficiency of traffic control in signalized road networks. Although the optimization of signal timings for an isolated junction is relatively easy, it requires further research due to "*offset*" and "*network cycle time*" components being in Coordinated Signalized Networks (CSN).

For the CSN, **TRA**ffic **N**etwork **S**tud**Y** **T**ool (TRANSYT) is one of the most useful tools for optimizing signal timings and also the most widely used program of its type. It is a stage-based optimization program and was developed by Transportation and Road Research Laboratory (Robertson, 1969). TRANSYT consists of two main parts: A traffic model and a signal timing optimizer. The traffic model utilizes a platoon dispersion model that simulates the normal dispersion of platoons as they travel downstream. It simulates traffic in a network of signalized intersections to produce a cyclic flow profile of arrivals at each intersection that is used to compute a Performance Index (*PI*) for a given signal timing and staging plan. The performance

* Corresponding author. Tel.: +90 258 2963351; fax: +90 258 2963460.
E-mail address: halimc@pau.edu.tr (H. Ceylan).

http://dx.doi.org/10.1016/j.trc.2015.03.010
0968-090X/

index is defined as the sum of a weighted linear combination of estimated delay and number of stops per unit time for all signal-controlled traffic streams and is used to measure the overall cost of traffic congestion associated with the traffic control plan.

TRANSYT version 7 was originally modified for the Federal HighWay Administration (FHWA), thus it was called "TRANSYT-7F." The *PI* in TRANSYT-7F may be defined in a number of ways. One of the TRANSYT-7F's *PI* is the Disutility Index (*DI*). The *DI* is a measure of disadvantageous operation; that is stops, delay, fuel consumption, etc. Optimization in TRANSYT-7F consists of a series of trial simulation runs using the simulation engine. Each simulation run is assigned a unique signal timing plan by the optimization processor. The optimizer applies the Hill-Climbing (HC) or Genetic Algorithm (GA) searching strategies. The trial simulation run resulting in the best performance is reported as optimal. Although the GA is mathematically better suited for determining global or near global optimal solution, relative to HC optimization, it generally requires longer CPU times than the HC optimization (McTrans Center, 2008).

In modeling the signal timing optimization problem, different objectives and methods have been sought in the literature. Wong (1995) provided approximate expressions for the derivatives of performance index with respect to the phase-based control variables. The derivatives calculated from these expressions were compared with those obtained from numerical differentiation. The results for both methods were found to be in good agreement but the approximate expressions required much less computational effort. Wong (1996) also proposed an approach for area traffic control using group-based control variables. In this case, the TRANSYT performance index is considered as a function of the group-based control variables, cycle time, start and duration of green time. About 10% improvements in the optimal performance was gained over the stage-based method in TRANSYT. Since this approach requires much longer computational time, parallel computing was also investigated to reduce the computational time for optimization of signal settings by Wong (1997). Heydecker (1996) proposed a decomposition approach to optimize signal timings based on group-based variables without taking the effect of the coordination between adjacent intersections into account. Afterwards, Wong et al. (2002) developed a time-dependent TRANSYT traffic model for the evaluation of performance index. Three scenarios have been considered and a microscopic simulation model has been used to evaluate the performance indices for the signal plans derived from these three scenarios. Using the proposed group-based methodology in Scenario 3, a remarkable improvement over Scenario 1 taking the average flows for analysis has been obtained. Moreover, when compared with the signal plans from Scenario 2 based on independent analyses, a good improvement has also been found. Girianna and Benekohal (2002) presented two different GA techniques which are applied on signal coordination for oversaturated networks. Their paper reveals that micro GA implementation on signal coordination problems reaches the near-optimal values of signal timing much earlier than simple GA implementation. Similarly, Ceylan (2006) developed a GA with TRANSYT-HC optimization tool, and proposed a method for decreasing the search space to solve the area traffic control problem. Proposed approach was found better than TRANSYT regarding optimal values of signal timings and performance index. Chen and Xu (2006) investigated the application of Particle Swarm Optimization (PSO) algorithm to solve signal timing optimization problem. Their results showed that PSO can be applied to this problem under different traffic demands. Dan and Xiaohong (2008) developed an improved GA in order to find optimal signal plans for signal optimization problem, which takes the coordination of signal timings for all signal-controlled junctions into account. The results showed that the method based on GA could minimize delay and improve capacity of network. Li (2011) presented an arterial signal optimization model that consider queue blockage among intersection lane groups under oversaturated conditions. The proposed model captures traffic dynamics with the cell transmission concept, which takes into account complex flow interactions among different lane groups. Liu and Chang (2011) further developed an arterial signal optimization model which considers physical queue evolution on arterial links by lane-group and the dynamic interactions of spillback queues among lane groups. The solution procedure developed with GA has been tested with an example arterial under different demand scenarios. Results revealed that the proposed model may be considered for use in design of arterial signals in comparison with TRANSYT-7F. He et al. (2012) presented a platoon-based mathematical formulation, which aims to provide multi-modal dynamical progression on the arterial based on the probe information, to perform arterial traffic signal control. VISSIM software shows that the proposed model can easily handle two common traffic modes, transit buses and automobiles, and significantly reduce delays for both modes under both non-saturated and oversaturated traffic conditions as compared with timings optimized by SYNCHRO. Jones et al. (2013) addressed the problem of determining robust signal controls in a road network considering interdependency of signal controls and traffic flow patterns and uncertainty in the travel demands. According to the results of case studies performed, their approach seems to provide a robust performance for solving signal control problem. On the other hand, Hu and Liu (2013) developed a data-driven arterial offset optimization model taking some inherent problems with vehicle-actuated signal coordination into consideration. The aim of this model is to minimize total delay for the main coordinated direction and to maximize the performance of the opposite direction as well. Results obtained from the field experiments show that the proposed model can reduce travel delay of coordinated direction significantly without compromising the performance of the opposite approach. Hu et al. (2013) proposed a model which maximizes the discharging capacity along oversaturated routes by considering green time constraints. In order to obtain the maximum flow, a forward–backward procedure was used in the model which tested using a microscopic traffic simulation model for an arterial network. Results indicated that the model can effectively reduce oversaturation and thus improve system performance. Varaiya (2013) considered the control of a network of signalized intersections and introduced the max pressure control which selects a stage that depends only on the queues adjacent to the intersection. Results show that max pressure control with some modifications which guarantees minimum green for each approach and considering of weighted queues is able to control of signalized networks although priority