

Link-based Parameterized Micro-tolling Scheme for Optimal Traffic Management

Extended Abstract

Hamid Mirzaei
University of California, Irvine
Irvine, CA
mirzaei@uci.edu

Guni Sharon
University of Texas at Austin
Austin, TX
gunisharon@gmail.com

Stephen Boyles
University of Texas at Austin
Austin, TX
sboyles@mail.utexas.edu

Tony Givargis
University of California, Irvine
Irvine, CA
givarigs@uci.edu

Peter Stone
University of Texas at Austin
Austin, TX
pstone@cs.utexas.edu

ABSTRACT

In the micro-tolling paradigm, different toll values are assigned to different links within a congestible traffic network. Self-interested agents then select minimal cost routes, where cost is a function of the travel time and tolls paid. A centralized system manager sets toll values with the objective of inducing a user equilibrium that maximizes the total utility over all agents. A recently proposed algorithm for computing such tolls, denoted Δ -tolling, was shown to yield up to 32% reduction in total travel time in simulated traffic scenarios compared to when there are no tolls. Δ -tolling includes two global parameters: β which is a proportionality parameter, and R which influences the rate of change of toll values across all links. This paper introduces a generalization of Δ -tolling which accounts for different β and R values on each link in the network. While this enhanced Δ -tolling algorithm requires setting significantly more parameters, we show that they can be tuned effectively via policy gradient reinforcement learning. Experimental results from several traffic scenarios indicate that Enhanced Δ -tolling reduces total travel time by up to 28% compared to the original Δ -tolling algorithm, and by up to 45% compared to not tolling.

KEYWORDS

Micro-tolling; Policy Gradient; Reinforcement Learning

ACM Reference Format:

Hamid Mirzaei, Guni Sharon, Stephen Boyles, Tony Givargis, and Peter Stone. 2018. Link-based Parameterized Micro-tolling Scheme for Optimal Traffic Management. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018*. IFAAMAS, 3 pages.

1 INTRODUCTION

Recent advances in connected vehicle technology enable new traffic management paradigms. For example, the notion of Micro-tolling, which has been the focus of some recently proposed tolling mechanisms [9–11], is that tolls can be assigned to all links in a given traffic network and changed depending on the traffic conditions in

real-time. Many methods have been proposed in the past to set toll values in order to achieve the optimal traffic flow. However, most of these methods are based on specific assumptions that might not hold in real life.

In contrast, Δ -tolling is a recently proposed scheme [10, 11] that makes no assumption on the underlying traffic model, e.g. fixed or known demand, link capacity, user value of time (VOT) or specific traffic formulation. In Δ -tolling, tolls are set proportional to the difference between the current travel time and the free flow travel time (denoted by Δ). The proportionality factor is denoted by β . Furthermore, the rate of change in toll values between successive time steps is regularized by another parameter R . While being simple to calculate and model free, Δ -tolling was shown to optimize the system's performance[12].

In this paper, we propose to generalize Δ -tolling by assigning different β and R parameters to each link of the network. This generalization comes with the cost of increased number of β and R parameters that should be tuned. Therefore, we also propose a tuning algorithm based on policy gradient reinforcement learning (RL). Our empirical studies show that the generalized algorithm considerably improves the performance of the tolling scheme compared to the original Δ -tolling.

2 PROBLEM DEFINITION AND BACKGROUND

This paper considers a scenario where there is a set of agents that should be routed across a traffic network. Agents are assumed to be self-interested, i.e., they choose the path with minimum cost. The cost of a path is the path's latency plus the tolls paid on that path. Since toll values might change over time, agents are continually looking for the optimal route and may change en route.

The system manager should assign tolls such that the total travel time is minimized while each agent maximizes its own self-interest. Formally, the micro-tolling assignment problem is defined as follows.

Given: link latencies measured at current time step i .

Output: the toll values for each link that should be applied at the next time step.

Objective: Minimize total system travel time.

Assumption: Agents are self-interested. This means that each

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden. © 2018 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

agent chooses the route that is at its best interest (minimal travel time).

The methods presented in this paper are based on two existing algorithms: Δ -tolling, and policy gradient RL. In the following subsections, each of these algorithms are described briefly.

2.1 Delta-tolling

It has been previously proven that charging each agent the marginal cost, i.e. the cost it inflicts on other agents, leads to optimal system performance [8]. However, calculating marginal cost in real-world scenarios is not feasible. In [10, 11], a model-free method, i.e. Δ -tolling is proposed to approximate marginal cost. In Δ -tolling a toll is calculated for each link at any time step by multiplying the difference of current travel time and the free-flow travel time (Δ) by a constant parameter β . The actual toll applied on each link is smoothed by another parameter R to remove transient toll value spikes, according to the following equation:

$$\tau^{i+1} = R(\beta\Delta) + (1 - R)\tau^i, \quad (1)$$

where τ^i is assigned toll at time step i .

2.2 Policy gradient RL

A well-known approach to learn a parameterized policy based on on-line data is to use general purpose Policy gradient RL method. There are multiple different methods to estimate the policy gradient [7]. We have chosen *Finite Difference Policy Gradient RL* (FD-PGRL) [4]. Unlike other methods that need within-episode rewards or the agent should learn the policy with no domain knowledge, FD-PGRL can leverage an existing policy with a reasonable performance while making small changes to the policy parameters in order to proceed towards the optimal policy and it only uses finite differences to estimate the policy gradient.

3 ENHANCED DELTA-TOLLING

In this section we present the Enhanced Δ -tolling ($E\Delta$ -tolling) which is an extensions to Δ -tolling approach introduced in 2.1. The original Δ -tolling has two global parameters R and β for the whole network. $E\Delta$ -tolling extends the parameter set by assigning different values of R , β or both for each link. Therefore, the number of parameters in $E\Delta$ -tolling can be up to $2|E|$ where $|E|$ is the number of links. The increased number of parameters requires a feasible tuning algorithm since the tuning cannot be done in a brute-force way. We have used FD-PGRL introduced in 2.2. To use FD-PGRL, $E\Delta$ -tolling is defined as a parameterized policy. The policy parameters are β , R or both assigned for each link of the network.

While it is suggested in [11] that there is a correlation between the R and β parameters, no conclusion was provided regarding this correlation. Therefore, we consider three variants of $E\Delta$ -tolling:

- $E\Delta$ -tolling $_{\beta}$ - this variant uses a global R parameter and link specific β parameters ($|E| + 1$ parameters in total).
- $E\Delta$ -tolling $_R$ - this variant uses a global β parameter and link specific R parameters ($|E| + 1$ parameters in total).
- $E\Delta$ -tolling $_{\beta,R}$ - this variant uses link specific β and R parameters ($2|E|$ parameters in total).

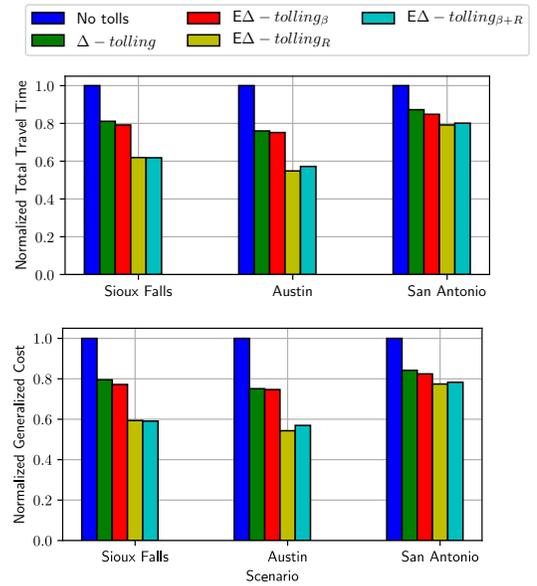


Figure 1: Total Travel Time and Total Generalized Cost for different tolling schemes and scenarios.

	Sioux Falls	Austin	San Antonio
Latency (hr)	11,859	21,590	26,362
cost (\$)	353,169	637,086	780,739

Table 1: Average total latency and total generalized cost when applying no tolls.

4 EMPIRICAL STUDY AND RESULTS

In all the experiments, traffic is modeled using the cell transmission model (CTM) [2, 3]. The DTA simulator [1] was used to run CTM and the simulation settings are the same as reported in [11]. Three traffic scenarios were used: Sioux Falls [5], Downtown Austin [6] and Uptown San Antonio which are available at: <https://goo.gl/SyvV5m>. The FD-PGRL parameters, i.e., step size, perturbation and number of policy runs at each step were set to 0.4, 0.01 and 60 respectively.

Total latency over all trips and total travel cost over all agents (social welfare) are presented in Figure 1. The values are normalized with respect to the no-toll scenario. No-toll scenario performance is summarized in Table 1.

The results suggest that tuning R parameter per link while having a global β leads to the best system performance in most cases.

REFERENCES

- [1] Yi-Chang Chiu, Jon Bottom, Michael Mahut, Alex Paz, Ramachandran Balakrishna, Travis Waller, and Jim Hicks. 2011. Dynamic traffic assignment: A primer. *Transportation Research E-Circular E-C153* (2011).
- [2] Carlos F. Daganzo. 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B* 28, 4 (1994), 269–287.
- [3] Carlos F. Daganzo. 1995. The cell transmission model, part II: network traffic. *Transportation Research Part B* 29, 2 (1995), 79–93.
- [4] Nate Kohl and Peter Stone. 2004. Policy Gradient Reinforcement Learning for Fast Quadrupedal Locomotion. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation, ICRA 2004, April 26 - May 1, 2004, New Orleans, LA, USA*. 2619–2624. <https://doi.org/10.1109/ROBOT.2004.1307456>

- [5] L. J. LeBlanc, E. K. Morlok, and W. P. Pierskalla. 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation Research* 9, 5 (1975), 309–318.
- [6] Michael W Levin, Matt Pool, Travis Owens, Natalia Ruiz Juri, and S Travis Waller. 2015. Improving the convergence of simulation-based dynamic traffic assignment methodologies. *Networks and Spatial Economics* 15, 3 (2015), 655–676.
- [7] Jan Peters and Stefan Schaal. 2006. Policy Gradient Methods for Robotics. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2006, October 9-15, 2006, Beijing, China*. 2219–2225. <https://doi.org/10.1109/IROS.2006.282564>
- [8] Arthur Cecil Pigou. 1920. *The Economics of Welfare*. Palgrave Macmillan.
- [9] Guni Sharon, Michael Albert, Stephen Boyles Tarun Rambha, and Peter Stone. 2018. Traffic Optimization For a Mixture of Self-interested and Compliant Agents. In *32th AAAI Conference on Artificial Intelligence (AAAI-18)*.
- [10] Guni Sharon, Josiah P. Hanna, Tarun Rambha, Michael W. Levin, Michael Albert, Stephen D. Boyles, and Peter Stone. 2017. Real-time Adaptive Tolling Scheme for Optimized Social Welfare in Traffic Networks. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2017)*.
- [11] Guni Sharon, Michael W. Levin, Josiah P. Hanna, Tarun Rambha, Stephen D. Boyles, and Peter Stone. 2017. Network-wide adaptive tolling for connected and automated vehicles. *Transportation Research Part C* 84 (September 2017), 142–157. <https://doi.org/10.1016/j.trc.2017.08.019>
- [12] Yosef Sheffi. 1985. Urban transportation network. *Equilibrium analysis with mathematical programming methods*, Prentice Hall (1985).