In *Proceedings of the 9th International Workshop on Agents in Traffic and Transportation* (ATT 2016), New-York, NY, USA July 2016

Delta-Tolling: Adaptive Tolling for Optimizing Traffic Throughput

Guni Sharon[1], Josiah Hanna[1], Tarun Rambha[2], Michael Albert[1], Peter Stone[1], Stephen D. Boyles[2]
[1]Department of Computer Science, [2]Department of Civil Engineering
The University of Texas at Austin,
Austin, TX 78712 USA

Abstract

In recent years, the automotive industry has been rapidly advancing toward connected vehicles with higher degrees of autonomous capabilities. This trend opens up many new possibilities for AI-based efficient traffic management. This paper investigates traffic optimization through the setting and broadcasting of dynamic and adaptive tolls under the assumption that the cars will be able to continually reoptimize their paths as tolls change.

Previous work has studied tolling policies that result in optimal traffic flow and several traffic models were developed to compute such tolls. Unfortunately, applying these models in practice is infeasible due to the dynamically changing nature of typical traffic networks. Moreover, this paper shows that previously developed tolling models that were proven to yield optimal flow in theory may not be optimal in real-life simulation. Next, this paper introduces an efficient tolling scheme, denoted Δ-*tolling*, for setting dynamic and adaptive tolls. We evaluate the performance of Δ-tolling using a traffic micro-simulator. Δ-tolling is shown to reduce average travel time by up to 35% over using no tolls and by up to 17% when compared to the current state-of-the-art tolling scheme.

1 Introduction

In recent years, communication and computation capabilities have become increasingly common onboard vehicles. Such capabilities present opportunities for developing safer, cleaner and more efficient road networks. This paper combines knowledge from mechanism design, game theory, network flow optimization, and multi-agent simulations for investigating responsive pricing, as a scheme for managing and optimizing traffic flow.

It has been known for nearly a century that drivers seeking to minimize their private travel times need not minimize the total level of congestion. In other words, self-interested drivers may reach an equilibrium that is not optimal from a system perspective. On the other hand, charging each agent with an amount equivalent to the damage it inflicts on all other agents (also known as the *marginal cost*) results in optimal flow [Pigou, 1920; Beckmann *et al.*, 1956; Braess, 1969]. The damage inflicted by a given agent is evaluated through the marginal slowdown caused by it and is commonly evaluated using stylized traffic models. Such stylized models take a "macroscopic" view of traffic, where delay can be expressed as a smooth function of travel demand. We hereafter refer to such models as macro-models. The marginal slowdown, evaluated by such models, is then used to infer appropriate tolls. However these macro-models make many approximations and assumptions that don't hold in practice.

Modern simulation tools and computational power allow for much more fine-grained simulation of traffic networks, referred to as micro-simulation models. Using such a realistic traffic simulator we demonstrate the potential of using tolls for reducing average travel time and increasing average utility. In this paper we show (empirically) that computing tolls using a macro-model does not lead to optimal performance in a realistic simulator. We explain this effect by noting that macro-models assume deterministic conditions, and have a number of unrealistic features. In recent years, researchers have relaxed the assumptions of the first macroscopic tolling models to incorporate responsiveness to roadway disruptions such as accidents [de Palma and Lindsey, 1998; Yang, 1999a,b; Lindsey, 2009; Boyles *et al.*, 2010] and to the total level of travel demand [Nagae and Akamatsu, 2006; Chen and Subprasom, 2007; Gardner *et al.*, 2008, 2011]. However, the effectiveness of all of these models is still restricted by the use of simplifying assumptions such as constant and known demand and capacity for each link.

In response to the suboptimal performance of existing macro-models, this paper introduces a novel tolling scheme denoted Δ-tolling. Δ-tolling approximates the marginal cost of each link using only two variables (current travel time and free flow travel time) and one parameter. Due to its simplicity, Δ-tolling is fast to compute, adaptive to current traffic, and accurate. We prove that, under some assumptions, Δ-tolling results in tolls that are equivalent to the marginal cost and demonstrate that it can lead to near-optimal performance in practice.

2 Motivation

This section defines the notion of *user equilibrium* (UE) and *system optimum* (SO). Applying tolls is then introduced as a mechanism that allows UE and SO to coincide. The *marginal cost toll* (MCT) policy is then presented followed by some macroscopic traffic models that approximate it. We discuss some of the drawbacks of such macro-models, which provides the motivation for the current study.

2.1 Computing User Equilibrium

Consider a directed network $G = (V, E)$, where V and E are the set of nodes and links respectively. Suppose that the demand (flow rates) between every pair of nodes is known. In this paper we assume that the travel time on a link $e \in E$ is a function of its flow (x_e) and is represented using a non-decreasing function $t_e(x_e)$ (also called volume delay or link-performance functions). In practice, the Bureau of Public Roads (BPR) function $t_e(x_e) = T_e(1 + \alpha(\frac{x_e}{C_e})^\beta)$ is commonly used as the delay function, where T_e is the free flow travel time and C_e is the capacity of link e. α and β are parameters whose default values are 0.15 and 4 respectively.

When agents choose routes selfishly, a state of equilibrium, called user equilibrium (UE) [Wardrop, 1952], is reached in which all used routes between an origin-destination (OD) pair have equal and minimal travel time. The link flow rates corresponding to this state can be obtained by solving a non-linear convex program that minimizes the Beckmann potential function ($\sum_{e\in E} \int_0^{x_e} t_e(x_e)\, dx$) Beckmann *et al.* [1956]. This objective ensures that the KKT (Karush-Kuhn-Tucker) conditions [Kuhn and Tucker, 1951; Karush, 1939] of the convex program correspond to Wardrop's UE principle [Wardrop, 1952]. The constraints of the optimization problem include non-negativity and flow conservation constraints. This model, also known as the traffic assignment problem (see Patriksson [1994] for a thorough overview), has been widely studied because of the mathematically appealing properties associated with convex programming.

2.2 Computing System Optimum

The system optimal (SO) problem can be formulated using a set of constraints similar to those used for computing UE but replacing the objective function with $\sum_{e\in E} x_e t_e(x_e)$. As mentioned before, all agents do not experience equal and minimal travel times at the SO state which incentivizes agents to switch routes. Instead, if an optimal tolling policy is applied, the flows resulting from a UE assignment in which agents minimize the generalized cost (time + toll) coincides with the SO solution. MCT is proven to be such a policy (UE=SO) [Pigou, 1920; Beckmann *et al.*, 1956; Braess, 1969]. In MCT each agent is charged a toll that is equal to the increase in travel time it inflicts on all other agents. Unfortunately, knowing in advance the marginal impact of an agent on traffic is infeasible in practice.

2.3 Approximating Marginal-Cost Tolls

The focus of this paper is methods that approximate the marginal cost. Most of these methods assume that demand on each link is constant. In such cases MCT can be formally defined as follows: given a link (e) and flow (x_e) the toll applied to e equals the change in travel time caused by an infinitesimal flow ($\frac{dt_e(x_e)}{dx_e}$) multiplied by the number of agents currently on this link (x_e).

A number of researchers have attempted to develop macro-models that approximate MCT for a given system [Yang *et al.*, 2004; Han and Yang, 2009]. However, a major drawback of such macro-models is that they are static and do not capture the time-varying nature of traffic. They also assume that the delay on each link is a function of its flow and hence neglect effects of intersections and traffic shocks. Although there has been some research on congestion pricing using finer traffic flow models, most of the existing models either assume complete knowledge of demand distribution over time [Wie and Tobin, 1998; Joksimovic *et al.*, 2005] or are restricted to finding tolls on freeways in which travelers choose only between parallel tolled and free general-purpose lanes [Gardner *et al.*, 2013, 2015; Yin and Lou, 2009]. This limitation motivates us to employ a simulation framework to replicate traffic in a more realistic manner, evaluate the performance of existing macro-models, and develop new methods to determine optimal tolls while adapting to unknown and changing demand.

3 Simulation

In order to evaluate the effectiveness of different tolling models on traffic flow optimization, we used a modified version of the *Autonomous Intersection Manager* (AIM) micro-simulator [Dresner and Stone, 2008]. On the one hand, AIM is very realistic in the sense that it allows simulating accelerations of individual vehicles in response to traffic conditions. On the other hand, due to computational limitations, AIM cannot scale to large road networks (only up to 3×3 grid network). For our experiments AIM was chosen since, unlike other simulators, it allows non deterministic traffic behavior, provides (direct) measurements on vehicle following distances, lane changes, gap acceptance, etc.

3.1 Autonomous Intersection Manager Simulator

AIM provides a multiagent framework for simulating autonomous vehicles on a road network grid; it presents a realistic traffic flow model that allows experimenting with adaptive tolling. The AIM simulator uses two types of agents: intersection managers, one per intersection, and driver agents, one per vehicle. Intersection managers are responsible for directing the vehicles through the intersections, while the driver agents are responsible for controlling the vehicles to which they are assigned. To improve the throughput and efficiency of the system, the driver agents "call ahead" to the intersection manager and request a path reservation (space-time sequence) within the intersection. The intersection manager then determines whether or not this request can be met. If the intersection manager approves a driver agent's request, the driver agent must follow the assigned path through the intersection. On the other hand, if the intersection manager rejects a driver agent's request, the driver agent may not pass through the intersection but may attempt to request a new reservation.

At every intersection, the driver agent navigator runs an A^* search [Hart *et al.*, 1968] to determine the shortest path leading to the destination of the vehicle associated with it. The navigator then directs the driver agent to drive via the shortest route. This behavior ensures that each vehicle acts greedily with respect to minimizing travel time. Next, we describe the required enhancements to the standard AIM simulator [Dresner and Stone, 2008] necessary to simulate realistic tolling experiments.

3.2 Enhancements to the AIM Simulator

In order to evaluate adaptive-tolling using AIM the following modifications were required:

- **Link toll:** each link (e) in the road network is associated with a toll, $toll_e$, which can adapt in real-time according to traffic conditions.
- **Link travel time:** each link stores: (1) an estimated travel time, t_e, that is based on real-time observed flow speed, and (2) an estimated free flow travel time T_e, that is based on the link's length divided by its speed limit.
- **Route selection:** when a car has several routes leading to its destination, the driver agent chooses the route ($r = e_1, e_2, ..., e_3$) that minimizes $\sum_{e \in r} t_e \times VOT + toll_e$, where VOT is the monetary *value Of time*.
- **Value Of Time:** each driver agent is associated with a randomly generated VOT that is drawn from a normal distribution. We assume monetary units are chosen such that the mean value is 1¢ per second, and assume a standard deviation of 0.2. VOT represents the value (in cents) of one second for the driver. A driver with $VOT = x$ is willing to pay up to x¢ in order to reduce travel time by 1 second.

3.3 Macroscopic Model

This paper uses a macroscopic model to approximate MCT. This model is used to solve the convex program described in Section 2 using Algorithm B [Dial, 2006]. Algorithm B is a bush-based/origin-based algorithm which exploits the fact that at equilibrium, all used routes carrying demand from a particular origin must belong to an acyclic subgraph in which each destination can be reached from the origin (such a subgraph is also called a bush). At each iteration, the algorithm maintains a collection of bushes (one for each origin), shifts agents within a bush to minimize their generalized costs, and adds/removes links in a bush until equilibrium is reached. Closeness to equilibrium is measured using *average excess cost*, which represents the average of the difference between each agent's generalized cost and the least cost path at the current flow solution. In the experiments presented in this paper, the algorithm was terminated when the average excess cost of the flow solution dropped below 1E-13.

4 Empirical Evaluation: Macroscopic Model

One of the main contributions of this paper is an empirical demonstration that setting tolls based on macro-models can lead to suboptimal results when evaluated in a more realistic micro-simulator. This section presents these empirical results, which motivate our new tolling scheme as presented in the next section.

4.1 Exemplar Road-Network

Figure 1 illustrates an exemplar road network that demonstrate the impact of tolls that adapt to traffic demand. The

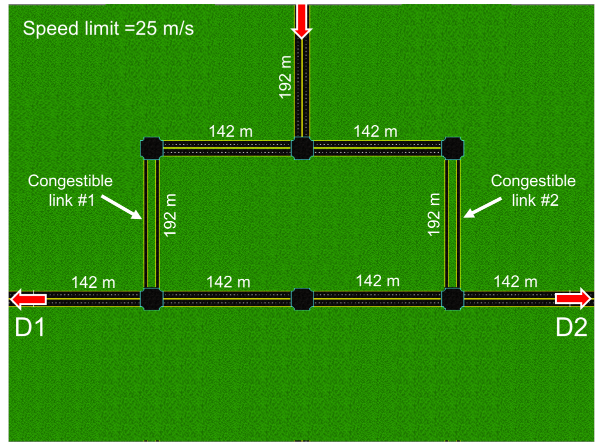

Figure 1: Exemplar road network within the AIM simulator.

speed limit across all roads is 25 meters per second. Each horizontal road is 142 meters long, and each vertical road is 192 meters long. We examined a scenario in which agents enter the network from a single source, the top road (incoming arrow), and leave the network from one of two destinations (outgoing arrows) D1 or D2. All roads are composed of two lanes per direction and assumed to have infinite capacity[1] except the two vertical roads in the middle of the network (Congestible link #1 and #2), which possess only one lane (capacity = 1,908 agents per hour). An agent entering the system and heading towards D1 (or symmetrically D2) has two possible routes to choose from: a short route (668 m) or a long route (964 m). Each agent chooses one of the two routes according to the distance, traffic conditions, and tolls associated with it. This road network represents a special case where if most agents are heading to D1 (or symmetrically D2) then link #1 (#2) should be tolled while link #2 (#1) should not. We define z (or symmetrically $1 - z$) to be the proportion of agents heading to $D1$ ($D2$). The incoming traffic rate was set to 2,160 agents per lane per hour.

4.2 Computing the Optimal Tolls

First, we computed, in a brute-force manner, the toll values that optimize average travel time for each $z \in \{0.0, 0.1, 0.2, ..., 1\}$. We considered tolling only congestible link #2. Tolling uncongestible links is unnecessary as there is no congestion externality associated with travel on these links. Moreover, there is no reason to toll both congestible links simultaneously (#1 and #2) since any of the two possible routes (leading from source to Di) includes exactly one of these links. A negative toll value for link #2 is symmetrical to a positive toll on link #1. We distinguish between the optimal *adaptive toll* and the optimal *static toll*. The optimal adaptive toll is the toll value that minimizes travel time for a given z value. The optimal static toll is the toll value that minimizes travel time over all z values (assuming equal weighting of the z values, i.e., all z values have the same probability), found to

[1] The capacity on roads with two lanes is higher than the rate in which agents are spawned. Hence, we consider such roads as having infinite capacity.

	Toll Values		AVG Travel Time (seconds)				
z	Optimal	Macro Model	No Tolls	Static Tolls	Optimal Tolls	Macro Tolls	Δ-Tolls
0.0	15	14.8	46.0	47.6	40.9*	**40.3***	40.5*
0.1	10	14.8	43.2	45.1	39.1*	39.3*	**39.0***
0.2	10	14.8	38.4	40.8	**35.8***	38.4	36.9*
0.3	10	14.8	34.3	35.1	33.8	37.7	**33.1***
0.4	0	14.8	31.7	32.4	31.7	36.8	**31.4**
0.5	5	-5.3	**30.8**	31.2	**30.8**	30.9	30.9
0.6	5	-14.8	**31.1**	31.5	**31.1**	34.7	31.6
0.7	-5	-14.8	**32.2**	32.2	**32.2**	35.2	32.8
0.8	-10	-14.8	37.0	**34.1***	**34.1***	36.2	35.8*
0.9	-10	-14.8	40.7	**36.2***	**36.2***	36.8*	36.5*
1.0	-15	-14.8	43.1	39.0*	38.5*	**38.1***	38.7*

Table 1: The left side of the table shows the empirical optimal and macro-model predicted toll values (imposed on link #2) for different z values. The right side shows average travel times when no tolls, static tolls, optimal tolls, macro-model tolls and Δ-tolls are applied as calculated by the AIM simulator. * indicates statistical significance over no tolls (using unpaired t-test with $p_{value} = 0.05$).

be -10 in this example. While it might seem like the optimal static toll should be zero, asymmetries in the model arising from differences between left and right turns affect junction delays and skew the optimal static toll to one side.

Optimal adaptive tolls for each z value are presented in Table 1. Notice that as the z value increases, the optimal toll steadily decreases. Intuitively, when all agents go to one destination ($z = 0$ or $z = 1$) we need more of them to choose the longer route to achieve the optimal system flow, thus requiring a more extreme toll. When $z \approx 0.5$, a zero toll is optimal since agents that choose their longer route will only make congestion worse for agents going to the other destination. As a result, enforcing tolls for $0.2 < z < 0.8$ did not result in a significant improvement over enforcing no tolls. The reason that Table 1 present values different than zero for that range stems from noise and asymmetries in the model.

4.3 Evaluating Optimal Tolls Using a Macro-Model

We compared the empirically optimal tolls against the toll values predicted by the macro-model. Toll values calculated by the macro-model are also presented in Table 1. Table 1 also presents average travel time under different tolling policies (for now ignore the Δ-tolls column). Though the macro-model obtains near optimal results for the extreme z values and $z = 0.5$, it is sub-optimal for intermediate values. One explanation for this phenomenon is that the stylized congestion models assume that delays on a link are a function solely of flow on that link, ignoring interactions between links at intersections. For the extreme z values this assumption is more reasonable because almost all agents on congestible links are heading in the same direction. However for the intermediate values (excluding 0.5) the agents on the congestible links encounter traffic on the bottom horizontal link (by cars taking the longer route) causing changes in the capacity of the congestible links that cannot be captured by a stylized model. These results lead us to the following conclusions:

1. Tolls can reduce average travel time by up to 11% compared to applying no tolls (see $z = 0$).
2. Static tolls might have a negative effect in some cases (see $z < 0.6$).
3. The macro-model fails to achieve system optimal in some cases reaching up to 10% suboptimality (see $z = 0.3$).

Both static and adaptive macro-model based tolls (a) result in suboptimal performance and (b) require that the demand over all OD pairs is known and fixed. As a result, neither is applicable to real-world traffic. There is thus, a need for a new tolling scheme that is dynamic, adaptive, and results in near-optimal traffic flow.

5 Delta-tolling

This section introduces the main technical contribution of the paper, a new tolling scheme denoted Δ-tolling. Unlike macroscopic models, Δ-tolling is adaptive to unknown and changing link capacities and demands. We first define Δ-tolling and then prove, under mild assumptions, that it is equivalent to MCT.

Δ-tolling is defined over a directed network $G = (V, E)$ (a road network for example) with a set of current flow values (traffic volume for example). The output of Δ-tolling is a set of toll values, one toll value per link. We use t_e to denote the current flow time on link $e \in E$. Recall that T_e denotes the free flow travel time and $toll_e$ to denote the toll value assigned to link e. For each link (e), Δ-tolling assigns a toll equivalent to the difference between the current flow time (t_e) and the free flow time (T_e) multiplied by a parameter (β). More formally: $\Delta\text{-}toll_e = \beta(t_e - T_e)$. As a rule of thumb, β should be correlated to the mean VOT. High β values result in high toll values which are needed to influence agents with high VOT. Calculating $\Delta\text{-}toll_e$ requires a constant amount of time. As a result, the complexity of computing tolls for an entire network is $\Theta(E)$.

Next we prove that Δ-tolling is equivalent to MCT under some conditions. This is a desirable property, since MCT results in system optimal (see Section 2). First, we list the assumptions under which the above statement holds:

	Macro-model	Δ-tolling
Parameters	Required	
α	yes	**no**
β	yes	yes
Variables	Required	
Demands	yes	**no**
C_e	yes	**no**
T_e	yes	yes
t_e	**no**	yes
Properties	Satisfied	
Adaptive	no	**yes**
MCT	**yes**	**yes**

Table 2: The different parameters, variables and properties of Δ-tolling and macro-model tolling. MCT refers to approximating the marginal cost.

1. The delay on each link is expressed by the BPR volume delay function, $t_e(x_e) = T_e(1 + \alpha(\frac{x_e}{C_e})^\beta)$.
2. Changes in traffic flow are negligible between the time an agent plans its route and the time it traverses the network.

Lemma 1 *Under the above assumptions, the tolls computed by Δ-tolling are equivalent to the MCT.*

Proof: We express the BPR volume delay function as:
(1) $t_e(x_e) = T_e + a{x_e}^\beta$ where $a = T_e\frac{\alpha}{{C_e}^\beta}$. MCT is defined as the derivative of the delay function ($\frac{dt_e(x_e)}{dx_e}$) multiplied by the flow (x_e). Calculating MCT requires knowing the future flow but under Assumption 2 we can use current flow instead. So we get:
(2) $MCT_e = x_e\frac{dt_e(x_e)}{dx_e} = x_e(\beta a{x_e}^{\beta-1}) = \beta a{x_e}^\beta = \beta(T_e + a{x_e}^\beta - T_e)$.
Combining (1) and (2) we get:
$MCT_e = \beta(t_e - T_e) = \Delta\text{-}toll_e$. □

The main theoretical differences between Δ-tolling and macroscopic models are summarized in Table 2. In the next section we study the differences in performance using the adapted AIM simulator.

Although the assumptions made in this section might not hold in all possible traffic networks, we provide experimental results showing that in realistic simulations, Δ-tolling improves traffic flow and may achieve near optimal flow.

6 Empirical Evaluation: Delta-Tolling

This section analyzes the performance of Δ-tolling on a representative road network. We then generalize our findings and show they also hold for randomly generated networks. We begin by comparing the system performance when using Δ-tolling on the exemplar road network (presented in Figure 1). Table 1 also presents the average travel time for Δ-tolling. Unlike the macro-model, Δ-tolling achieves performance that is similar to the optimal. We do not report the toll values for Δ-tolling as they are dynamically changing across the simulation.

Next, we present results for larger networks. In such networks finding the optimal tolls in a brute force manner is infeasible.[2] For the following experiments we used grid networks of size 3×3 that include 9 intersection (see Figure 3 for an example). Agents enter at the same rate of 300 agents per hour from any incoming lane (a road with three lanes, for example, spawns 900 agents per hour). Each agent entering the system is assigned one of two possible exit roads with equal probability (0.5). Each agent is also assigned two alternative exits. Exiting via an alternative exit imposes a predefined, randomly generated, delay.[3] We justify allowing alternative exits as follows, in many real-life scenarios, several routes, usually of different length, may lead an agent to its destination. For example, a driver exiting *Manhattan* and heading to *Queens* will prefer to exit via *Queens Midtown Tunnel*, it can suffer some delay and exit from *Ed Koch Queensboro Bridge* or suffer a severe delay while exiting via *Williamsburg Bridge*. Following this logic, we view the simulated network as part of a larger road network in which agents may use paths outside of the network to reach their final destination.

Some roads in the simulated network are more congestible than others i.e., the number of lanes varies. The number of lanes for each road was randomly picked from $[1, 4]$. We ran the simulator for 5000 seconds for each reported setting.[4] In the following experiments we used an upper bound on toll values equal to 25¢.[5] The upper bound is set for two reasons: (1) avoiding overcharging in links with temporary heavy congestion (2) avoiding oscillation in congestion caused by overpricing: heavy congestion may cause a steep increase to the toll value which later leads to the link being vacated which leads the toll value to reduce to zero. Zero toll value results, again, in heavy congestion. Applying no cap on toll values resulted in up to 5% reduced utility. We report three different measurements:

- **Time** - the average travel time.
- **Utility** - the average utility (in cents). Where utility is defined for each agent as its travel time multiplied by its VOT plus the summation of tolls paid by it.
- **Standardized Utility (SU)** - toll revenue may be redistributed back to the drivers in the form of road improvements or tax reductions. We define *refund* as the sum of collected tolls divided by the number of agents that exited the system. SU is defined as average utility minus refund.

6.1 Representative Road Network

The purpose of our first experiment is to determine how different β values affect system performance. For this experiment we used a single randomly generated instance of a 3×3

[2] Examining different combinations of toll assignment to all links in the system leads to an exponential blowup.

[3] When each agent is assigned only one possible exit, distributing traffic becomes impossible in many cases. For such scenarios, imposing tolls did not have a positive effect in our experiments.

[4] When running the simulator, in order to allow the system to balance, we exclude data from the first 500 seconds.

[5] The output from the macro-model contained no toll greater than 25¢. Hence we deduced that greater tolls won't have a positive affect and we set the cap accordingly.

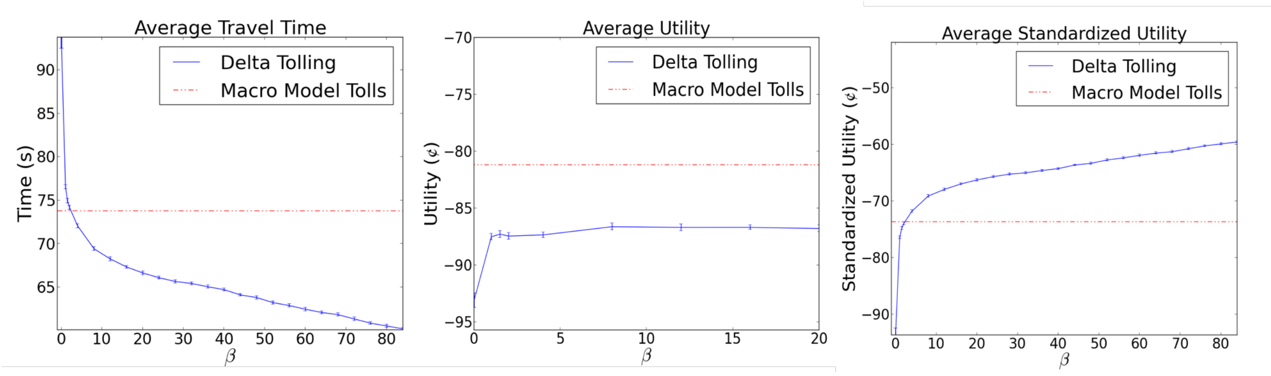

Figure 2: Average travel time, utility and standardized utility as a function of β for the representative road network.

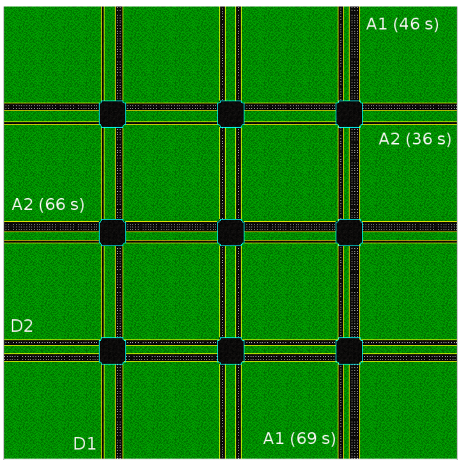

Figure 3: A representative road network. Each agent is assigned one of to destinations (**D1, D2**). **A1** and **A2** denote alternative destinations for **D1** and **D2** respectively. The time penalty associated with each alternative destination is given in parenthesis.

road network - depicted in Figure 3. Average travel time, Utility and SU for different β values are presented in Figure 2. Notice that $\beta = 0$ represent the case where no tolls are used.

Setting $\beta = 80$ gives an improvement of 35% in average travel time over no tolls. $\beta = 80$ also gives an improvement of 35.01% for SU over no tolls. β values greater than 80 result in average travel times that are not significantly worse or better. Increasing β (up to 80) results in higher toll values which better distribute congestion. However, higher tolls also negatively impact utility as drivers are forced to pay more. Utility is maximized with $\beta = 8$ which gives a 6.96% improvement over no tolls. We also report performance when tolls as computed by the macro-model are used, given as a dashed (red) line across the result graphs. Δ-tolling outperforms macro-model tolling for $\beta \geq 4$ by up to 18% in both average travel time and SU. On the other hand, macro-model tolling exceeds by 6.25% when utility is considered. The main reason for the macro-model's advantage w.r.t utility is that Δ-tolling imposes higher toll values. Δ-tolling (with $\beta = 8$) collected a total of \$1,921 while macro-model tolling collected only \$825. Unfortunately, we observed that higher tolls are required to better distribute congestion and optimize system performance. On the other hand, we believe that standard utility is a more relevant measurement for performance comparison between the models. In real road networks tolls are most often used to fund road maintenance, effectively redistributing the money collected back to the public. When SU is considered, delta tolling significantly outperforms the macro-model in all but very low β. Moreover, macro-model tolling relies on static traffic conditions and so, unlike Δ-tolling it is not applicable to real-life, dynamic road networks.

β	Time	Utility	SU
0.0	**69.9**	**-70.0**	**-70.0**
8.0	51.4*	-63.5	-51.1*
20.0	50.3*	-67.0	-49.8*
80.0	49.5*	-76.6	-48.8*

Table 3: Average travel time, utility, and SU for β values $8, 20, 80$. These β values represent a trade-off between the three metrics. *Denotes a statistically significant improvement over no tolling (using a paired t-test with $p_{value} = 0.05$).

6.2 General Case

In order to validate that the results obtained from a single randomized instance are representative, we reran the same experiment using 50 different randomized road networks. Each of these networks is a 3×3 grid, similar to the representative road network, but the exit roads, alternative exits, alternative exits' delay, and number of lanes per road are randomized. Table 3 shows results for three representative β values (8, 20, 80) compared to no tolling. $\beta = 8$ and $\beta = 80$ are chosen since they maximized performance with respect to utility and travel time/SU. $\beta = 20$ represents a good balance between utility and travel time.

We observe that the advantage of Δ-tolling is robust to changes in network topology. For the general case, Δ-tolling achieves an improvement over no tolling of 29.2%, 9.31% and 30.28% in Time, Utility and SU respectively.

7 Conclusions

This paper considers applying tolls to road networks in order to direct the route choice of self-interested agents towards a system optimal. The notion of such a tolling scheme is becoming more practical as cars are becoming increasingly autonomous and the computational effort required to evaluate several alternative routes is becoming more feasible.

This paper makes two main contributions. First, using a traffic micro-simulator (AIM), we provide empirical evidence suggesting that stylized macroscopic traffic models are unable to approximate optimal tolls accurately. Given this finding and the fact that such models require full knowledge of demand and supply and assume that these remain fixed, we conclude that using such models to set tolls in real-life road networks is impractical. This conclusion leads us to the second contribution, the presentation and evaluation of a new tolling scheme, denoted Δ-tolling. We provide theoretical and empirical evidence that Δ-tolling results in near-optimal system performance while being adaptive to traffic conditions and computationally feasible.

Our ongoing research agenda includes evaluating the performance of Δ-tolling in dynamic environments, in which traffic demand and supply is time varying.

8 Acknowledgments

A portion of this work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CNS-1330072, CNS-1305287), ONR (21C184-01), and AFOSR (FA9550-14-1-0087). Peter Stone serves on the Board of Directors of, Cogitai, Inc. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research. The authors would also like to acknowledge the support of the Data-Supported Transportation Operations & Planning Center and the National Science Foundation under Grant No. 1254921.

References

M. J. Beckmann, C. B. McGuire, and C. B. Winston. *Studies in the Economics of Transportation*. Yale University Press, New Haven, CT, 1956.

Stephen D. Boyles, Kara M. Kockelman, and S. Travis Waller. Congestion pricing under operational, supply-side uncertainty. *Transportation Research Part C*, 18(4):519–535, 2010.

Dietrich Braess. Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1969.

Anthony Chen and Kitti Subprasom. Analysis of regulation and policy of private toll roads in a build-operate-transfer scheme under demand uncertainty. *Transportation Research Part A*, 41(6):537–558, 2007.

Andr'e de Palma and Robin Lindsey. Information and usage of congestible facilities under different pricing regimes. *Canadian Journal of Economics*, 31(3):666–692, 1998.

Robert B. Dial. A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration. *Transportation Research Part B*, 40(10):917–936, 2006.

Kurt Dresner and Peter Stone. A multiagent approach to autonomous intersection management. *Journal of artificial intelligence research*, pages 591–656, 2008.

Lauren Gardner, Jennifer Duthie, Avinash Unnikrishnan, and S. Travis Waller. Robust pricing for networks with demand uncertainty, 2008. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, DC.

Lauren Gardner, Stephen D. Boyles, and S. Travis Waller. Quantifying the benefit of responsive pricing and travel information in the stochastic congestion pricing problem. *Transportation Research Part A*, 45:204–218, 2011.

Lauren Gardner, Hillel Bar-Gera, and Stephen D. Boyles. Development and comparison of choice models and tolling schemes for high-occupancy/toll (HOT) facilities. *Transportation Research Part B*, 55:142–153, 2013.

Lauren Gardner, Stephen D. Boyles, Hillel Bar-Gera, and Kelly Tang. Robust tolling schemes for high-occupancy/toll (HOT) facilities under variable demand. *Transportation Research Record*, 2450:152–162, 2015.

Deren Han and Hai Yang. Congestion pricing in the absence of demand functions. *Transportation Research Part E: Logistics and Transportation Review*, 45(1):159 – 171, 2009.

Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*, 4(2):100–107, 1968.

Dusica Joksimovic, Michiel C.J. Bliemer, and Piet H.L. Bovy. Optimal toll design problem in dynamic traffic networks with joint route and departure time choice. *Transportation Research Record: Journal of the Transportation Research Board*, 1923(1):61–72, 2005.

William Karush. *Minima of functions of several variables with inequalities as side constraints*. PhD thesis, Masters thesis, Dept. of Mathematics, Univ. of Chicago, 1939.

H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492, Berkeley, Calif., 1951. University of California Press.

Robin Lindsey. Cost recovery from congestion tolls with stochastic capacity and demand. *Journal of Urban Economics*, 66(1):16–24, 2009.

T. Nagae and T. Akamatsu. Dynamic revenue management of a toll road project under transportation demand uncertainty. *Networks and Spatial Economics*, 6:345–357, 2006.

M. Patriksson. *The Traffic Assignment Problem — Models and Methods*. VSP, Utrecht, Netherlands, 1994.

Arthur C. Pigou. *The Economics of Welfare*. Macmillan and Co., London, 1920.

John G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers, Part II*, 1:352–362, 1952.

Byung-Wook Wie and Roger L. Tobin. Dynamic congestion pricing models for general traffic networks. *Transportation Research Part B: Methodological*, 32(5):313 – 327, 1998.

Hai Yang, Qiang Meng, and Der-Horng Lee. Trial-and-error implementation of marginal-cost pricing on networks in the absence of demand functions. *Transportation Research Part B: Methodological*, 38(6):477 – 493, 2004.

Hai Yang. Evaluating the benefits of a combined route guidance and road pricing system in a traffic network with recurrent congestion. *Transportation*, 26:299–322, 1999.

Hai Yang. System optimum, stochastic user equilibrium, and optimal link tolls. *Transportation Science*, 33(4):354–360, 1999.

Y. Yin and Y. Lou. Dynamic tolling strategies for managed lanes. *Journal of Transportation Engineering*, 135(2):45–52, 2009.