Applications of Dynamic Pricing in Day-to-Day Equilibrium Models

Tarun Rambha
Graduate Research Assistant
The University of Texas at Austin
Department of Civil, Architectural, and Environmental Engineering
1 University Station C1761
Austin, TX 78712
tarun.1988@gmail.com
Ph: (512) 761-2551
Fax: (512) 475-8744
(*Corresponding author*)

Stephen D. Boyles
Assistant Professor
The University of Texas at Austin
Department of Civil, Architectural, and Environmental Engineering
1 University Station C1761
Austin, TX 78712
sboyles@austin.utexas.edu
Ph: (512) 471-3548
Fax: (512) 475-8744

Word count:
Text: 4067 (abstract: 142)
Tables: $3 \times 250 = 750$
Figures: $3 \times 250 = 750$
Total: 5567

Submission date: November 14, 2014

ABSTRACT

The traffic assignment problem is primary concerned with the study of user equilibrium and system optimum states. These models, however, require that travelers are perfectly rational and have a complete knowledge of network conditions. For an empirical standpoint, when a large number of selfish travelers control the flow on a network, the chances of reaching equilibrium are slim. User behavior in such settings can be modeled using probabilistic route choice models, which define when and how players switch paths. In the context of the traffic assignment problem, only a few dynamic route adjustment processes exhibit asymptotic convergence to equilibrium. In this paper, we propose a Markov decision process formulation for improving the probability of convergence of any closed-form route choice model to an equilibrium solution using dynamic pricing. A simple example to illustrate the application of the pricing framework is also discussed.

1 INTRODUCTION

Urban transportation planning is traditionally carried out using a four-step process. The first three steps are used to estimate the number of travelers/users, their origin-destination (OD) pairs, and their mode of travel. The final step, also called route choice or traffic assignment, involves assigning travelers to different routes. This assignment procedure is done assuming that traffic networks are in a state of equilibrium due to selfish choices made by travelers *(1, 2)*. In order to observe a user equilibrium (UE) or Nash equilibrium (NE) state, travelers must be rational and have a perfect knowledge of the network topology and its response to congestion. However, when a large number of travelers interact, the extent of reasoning required to arrive at an equilibrium solution remains beyond one's human ability. Learning models in transportation and behavioral game theory have tried relaxing these assumptions of perfect rationality to develop dynamic day-to-day models that take into account how travelers' choices vary with time in the presence of historical information on network conditions. A major goal of these studies has been to test if a particular dynamical process converges to equilibrium. However, the number of dynamics that are known to converge to equilibrium in traffic networks are quite limited. This motivates us to raise the following question: "Irrespective of how traffic networks evolve, can we use a dynamic pricing/tolling mechanism to ensure the convergence of players' choices to a UE?"

While congestion pricing has been traditionally used to achieve social optimal (SO) flows, the purpose of pricing here is to guide users to an equilibrium state. The advantages of guiding players to reach an equilibrium are two-fold. First, in the presence of multiple equilibria, we may be able to guide players to an efficient one. Second, and more importantly, SO states are equivalent to UE states under modified link performance functions *(3)*. Therefore, reaching a SO state is as difficult as reaching an UE, and hence using dynamic pricing we might also be able to direct players to a SO state. The main hypothesis of this research is that for a given route switching mechanism, assuming that travelers make *more rational choices with time*, we can improve the probability of reaching (i) a UE by dynamically pricing the network for finite number of days or (ii) a SO by dynamically pricing the network for a finite number of days and then shift to a static pricing scheme using the marginal costs. Since UE states are self-enforcing, if users become experienced enough, revoke pricing would not alter the state of the network. On the other hand, a SO state is not self-enforcing, and hence we would have to continue to collect tolls using the marginal costs.

The rest of this paper is organized as follows. In Section 2, we summarize the literature on the topic of day-to-day dynamics. Section 3 discusses two commonly used modeling approaches that are used to study the evolution of traffic. In Section 4, we propose a Markov decision process (MDP) model for finding a dynamic pricing policy that improves the probability of convergence to an equilibrium solution and demonstrate it using a logit route choice model on a small network. Finally, in Section 5, we summarize our findings and discuss pointers for future research on this topic.

2 LITERATURE REVIEW

Several efficient algorithms exist for finding the UE solution to the traffic assignment problem (TAP) *(4, 5, 6, 7, 8, 9, 10)*. A separate line of research has investigated how an equilibrium might be reached in traffic networks. This issue has received considerable attention in the literature and equilibrium is modeled as a steady state of a stochastic process *(11, 12, 13, 14)*. Also called as day-to-day dynamic models, these stochastic processes result from randomness in users perceived travel times. These models can be placed in a larger context of learning in repeated and evolutionary games in which players' actions are chosen in response to the past history of a game. Some of the commonly used dynamics include best response mechanism *(15, 16)*, replicator dynamics *(17, 18)*, projection dynamics *(19)*, and Brown-von Neumann-Nash dynamic *(20)*. For a detailed description of these dynamics see *(21)*.

Only a few dynamics exhibit convergence to UE in traffic networks. Traffic can be modeled as congestion games which belong to the class of potential games that possess the finite improvement property *(22)*. According to this property, a dynamic in which, at each round of a repeated game, a single traveler switches paths so that he/she is strictly better off converges to a UE. Fictitious play process *(15, 16)* is another dynamic in which each traveler best responds to the empirical joint distribution of other traveler' actions. However, the actual equilibrium solution to which fictitious play process converges depends on factors such as initial beliefs, tie-breaking rules, and the manner in which beliefs are updated (i.e., sequential or simultaneous).

A more realistic class of learning models in game theory relaxes the assumption of perfect rationality. The central idea in these models is to define a stochastic process using the outcomes of a game as system states, and assume some dynamic which lets players move from one state to another. This process is modeled as a Markov chain and its stationary or steady state distribution is used to study equilibrium solutions. Travelers in these models are modeled using the concepts of inertia, myopic behavior, and mutations. Inertia suggests that travelers are unlikely to frequently switch paths. Myopic behavior implies that travelers choose actions to optimize their present travel times rather than discounted infinite-horizon travel times. Mutations reflect the assumption that travelers may "tremble" or make mistakes while choosing a path. Depending on the probabilities that are assigned to the strategies that are not best responses, different learning algorithms can be constructed *(23, 24, 25, 26, 27)*.

3 DISCRETE AND CONTINUOUS DAY-TO-DAY MODELS

In this section, we discuss two day-to-day models that are commonly found in literature. The first model is a discrete time Markov chain (DTMC) in which travelers follow the logit path choice model *(11)*. The second one is similar to the first but is defined in a continuous setting *(27)* and is hence modeled as a continuous time Markov chain (CTMC). Both models are assumed to have a finite number of travelers. These models will be illustrated using the following example, which will later be used for demonstrating the proposed dynamic pricing mechanism. Suppose two travelers wish to travel from node O_1 to D_1 (traveler 1) and from O_2 to D_2 (traveler 2) in the network shown in Figure 1. The link performance functions are indicated on the arcs. There are four different outcomes in this network. Suppose we denote the states/feasible flow solutions using the ordered pairs $(T, T), (B, B), (T, B)$, and (B, T), where T stands for top link (with link performance function x^2) and B for bottom link (with link performance function x). We will refer to these states using the numbers 1, 2, 3, and 4 respectively. The first and the second elements of the ordered pair represent the choices made by the traveler 1 and traveler 2 respectively. In these models, we will assume that each time a traveler makes a route choice he or she uses a probability distribution based on the logit path choice model while assuming that other travelers remain on their current paths. Note that while the link flow solution to this problem is unique, there are two path flow solutions (states 3 and 4).

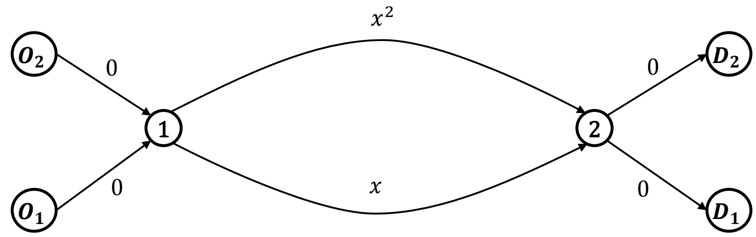

FIGURE 1: Example to demonstrate day-to-day dynamics in networks

3.1 Discrete time models

In the discrete version of day-to-day dynamic models, travelers make route choices on each day based on observed states in the past. Suppose travelers choices are conditioned on the current state of the system. This lets us define a Markov process with transition probabilities as defined below. Suppose $P = [p_{ij}]$ represents the transition matrix, where p_{ij} denotes the probability of moving from states i to state j in one time period.

$$
\begin{aligned}
p_{11} &= \text{Probability that both travelers stay on the top path} && (1)\\
&= \left(\frac{\exp(-4)}{\exp(-4)+\exp(-1)}\right)^2 && (2)\\
p_{12} &= \text{Probability that both travelers move to the bottom path} && (3)\\
&= \left(\frac{\exp(-1)}{\exp(-4)+\exp(-1)}\right)^2 && (4)\\
p_{13} &= \text{Probability that only traveler 2 switches to the bottom path} && (5)\\
&= \left(\frac{\exp(-1)}{\exp(-4)+\exp(-1)}\right)\left(\frac{\exp(-4)}{\exp(-4)+\exp(-1)}\right) && (6)\\
p_{14} &= \text{Probability that only traveler 1 switches to the bottom path} && (7)\\
&= \left(\frac{\exp(-1)}{\exp(-4)+\exp(-1)}\right)\left(\frac{\exp(-4)}{\exp(-4)+\exp(-1)}\right) && (8)
\end{aligned}
$$

The expressions for the other transition probabilities can be written similarly. The Markov chain defined using these transition probabilities is irreducible as every pair of states communicate with each other. Hence a steady state distribution exists. The long run percentage of finding the system in states 1, 2, 3, and 4 is 0.1885, 0.3201, 0.2456, and 0.2456 respectively. Since states 3 and 4 are the equilibrium states, one is likely to find the system in disequilibrium states nearly 50% of the time.

3.2 Continuous time models

In continuous time day-to-day models *(21, 27)*, all travelers do not change their paths at the same time but are presented with strategy revision opportunities at random intervals. The sojourn times for each traveler (time between two successive revision opportunities) are assumed to be exponentially distributed with rate λ. This framework is also called Logit learning and is modeled as a continuous time Markov chain (CTMC). When a traveler gets to choose a path (say at time t), he/she does so using a log-linear choice rule assuming that all other travelers remain on their current paths. Unlike in the case of discrete time models, players' transition rates are additionally assumed to be dependent on a parameter ϵ which defines the extent of making a mistake or the extent of irrationality. For any positive ϵ, the CTMC is irreducible and recurrent (since all states can communicate with each other). Hence, a unique steady state/limiting distribution that has all states in its support exists. However, as ϵ tends to zero, i.e., as the probability of making mistakes get smaller (it is assumed that by repeated interactions players get more experienced) only a few states have positive limiting probabilities. These states constitute what is termed a *stochastically stable set*. Blume *(27)* showed that as ϵ's tend to zero, the stochastically stable set coincides with the states that are in UE.

In the example discussed earlier, for a given ϵ, the transition probabilities for traveler 1, $\xi_1^\epsilon()$, can be written as shown below. When presented with a strategy revision opportunity, suppose traveler 2 is on the top path, traveler 1 chooses T and B with probabilities $\xi_1^\epsilon(T,T)$ and $\xi_1^\epsilon(B,T)$ respectively.

$$\xi_1^\epsilon(T,T) = \frac{\exp(-4/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \qquad \xi_1^\epsilon(B,T) = \frac{\exp(-1/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \tag{9}$$

$$\xi_1^\epsilon(T,B) = \frac{\exp(-1/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \qquad \xi_1^\epsilon(B,B) = \frac{\exp(-2/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \tag{10}$$

Expressions for the transition probabilities of traveler 2 may be written in a similar manner. The transition diagram and the associated transition rates are shown in Figure 2.

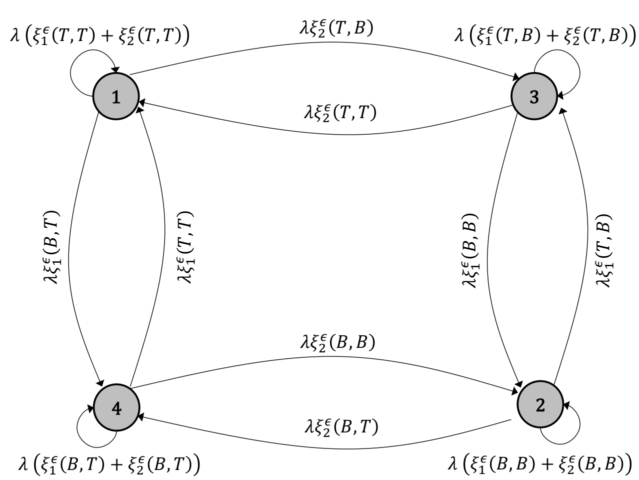

FIGURE 2: Transition diagram for a logit learning process

The global balance equations are then constructed and solved to obtain the long run proportion of time spent in each state.

$$\rho_1^\epsilon + \rho_2^\epsilon + \rho_3^\epsilon + \rho_4^\epsilon = 1 \tag{11}$$
$$\rho_1^\epsilon \left(\xi_2^\epsilon(T,B) + \xi_1^\epsilon(B,T)\right) = \rho_4^\epsilon \xi_1^\epsilon(T,T) + \rho_3^\epsilon \xi_2^\epsilon(T,T) \tag{12}$$
$$\rho_2^\epsilon \left(\xi_2^\epsilon(B,T) + \xi_1^\epsilon(T,B)\right) = \rho_4^\epsilon \xi_2^\epsilon(B,B) + \rho_3^\epsilon \xi_1^\epsilon(B,B) \tag{13}$$
$$\rho_3^\epsilon \left(\xi_2^\epsilon(T,T) + \xi_1^\epsilon(B,B)\right) = \rho_1^\epsilon \xi_2^\epsilon(T,B) + \rho_2^\epsilon \xi_1^\epsilon(T,B) \tag{14}$$
$$\rho_4^\epsilon \left(\xi_1^\epsilon(T,T) + \xi_2^\epsilon(B,B)\right) = \rho_1^\epsilon \xi_1^\epsilon(B,T) + \rho_2^\epsilon \xi_2^\epsilon(B,T) \tag{15}$$

Solving the balance equations for a given value of ϵ gives the steady state probabilities of finding the system in that state. As ϵ tends to zero, the support of the steady state probabilities is identical to set of UE solutions. Table 1 summarizes the behavior of the steady state probabilities for different values of ϵ. As can be seen from the table, the steady state probability of each of the UE states is 0.5 for low values of ϵ.

TABLE 1: Convergence of logit learning

ϵ	1	0.5	0.33	0.25	0.2	0.1	...	0.01
ρ_1^ϵ	0.020593	0.001159	6.02E-05	3.04E-06	1.52E-07	4.68E-14	...	2.6E-131
ρ_2^ϵ	0.152163	0.063305	0.024287	0.009075	0.003358	2.27E-05	...	1.86E-44
ρ_3^ϵ	0.413622	0.467768	0.487826	0.495461	0.498321	0.499989	...	0.5
ρ_4^ϵ	0.413622	0.467768	0.487826	0.495461	0.498321	0.499989	...	0.5

Alternately continuous time day-to-day dynamics can be modeled using ordinary differential equations *(28)* in which travelers are assumed to be infinitely divisible. Sandholm *(21)* shows that this approach is equivalent to the CTMC model when the number of travelers is large.

4 MODEL FORMULATION

We now propose a discrete time equivalent of the continuous time route choice model described in the previous section by including a term ϵ_k in the expressions for the transition probabilities, which represent the extent of irrationality on day k. We also suppose that travelers get more experienced over time and hence assume that $\epsilon_k \to 0$ as $k \to \infty$. However, unlike as seen in the continuous models, we demonstrate that the discrete time Markov chain with decreasing extent of irrationality can have steady state distributions in which a significant proportion of time is spent in disequilibrium states. We then propose a dynamic tolling framework that can potentially improve the probability of convergence to an equilibrium state. Further, in the presence of multiple equilibrium solutions, we will be able to choose the equilibrium to which we want the system to converge.

For the purpose of demonstration, suppose the ϵ's vary as $1/(k+1)$ and that the extent of irrationality is homogenous across all travelers. The transition probability matrix $P(k) = [p_{ij}(k)]$ is now a function of k. Given below is the modified version of (2), similar expressions for the other transition probabilities can be written.

$$p_{11} = \text{Probability that both travelers stay on the top path on day } k \tag{16}$$

$$= \left(\frac{\exp(-4/\epsilon_k)}{\exp(-4/\epsilon_k) + \exp(-1/\epsilon_k)}\right)^2 \tag{17}$$

Suppose that on day 0, all states are equally likely to be observed. The long run percentages of finding the system in each of the 4 states can be written as $a \lim_{k\to\infty} P(k)$, where a is a row vector that represents the initial distribution on day 0. Computing higher powers of P by direct matrix multiplication, the long run percentages was approximately found to be

$$= \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 0.0649634 & 0.555205 & 0.189916 & 0.189916 \\ 0.402912 & 0.133405 & 0.231841 & 0.231841 \\ 0.161785 & 0.272153 & 0.472968 & 0.0930937 \\ 0.161785 & 0.272153 & 0.0930937 & 0.472968 \end{bmatrix}$$
$$= \begin{bmatrix} 0.19786135 & 0.308229 & 0.246954675 & 0.246954675 \end{bmatrix} \tag{18}$$

As mentioned earlier, even if travelers get more experienced with time, the system is again likely to be found in disequilibrium states for nearly 50% of the time. In order to avoid this issue, we now propose a mechanism in which the links in the network are tolled for a finite number of days after which the tolls are revoked. Since users make more rational choices with time, once an equilibrium state is reached the probability of revisiting that state in the next time period is high, and hence the probability of finding the system in equilibrium can be improved. We assume that travelers value time and cost equally and hence attempt to minimize the sum of travel time and cost of their routes.

Consider N users who repeatedly make route choices between their origins and destinations over a finite number of days. We assume that the number of travelers are fixed. Although this assumption is limiting because there is usually a high degree of demand uncertainty in networks, the treatment of equilibrium with stochastic demand is a topic in itself and can be justly studied only if the problem with a fixed number of players is fully explored. Suppose that set of days is denoted by $\mathcal{K} = \{1, 2, \ldots, K\}$. Let $\mathbf{s}_k$ represent the vector of route choices on day k and $\mathbf{u}_k$ represent the vector of link tolls on day k. Assume that on any day, links in the network may be priced from a finite set of feasible tolls $\mathbf{U}$. We now define an MDP in which given the path choices on a particular day, one can determine the optimal toll that maximizes the probability (ideally one would want them to be close to 1) of reaching an equilibrium at the end of K^{th} day from every initial state. The tolls for day $k+1$ are chosen at the end of day k and are revealed to travelers before the beginning of the day $k+1$. The components of the MDP are as follows:

Time periods: $k = 0, 1, 2, \ldots, K$
States: Path choice vector on day k ($\mathbf{s}_k$). States are indexed by i, j
Actions: Tolls on day k ($\mathbf{u}_k$)
Transition probabilities: $p_{ij}^k(\mathbf{u}_k, \boldsymbol{\epsilon_k}) = \mathbb{P}[\mathbf{s}_{k+1} = j | \mathbf{s}_k = i, \mathbf{u}_k, \boldsymbol{\epsilon_k}]$
Value Functions: Probability of reaching a UE at the end of the time period ($J_k(\mathbf{s}_k)$)

Boundary conditions: $J_K(\mathbf{s}_K)$ is 1 if $\mathbf{s}_K$ is a UE and is 0 otherwise.

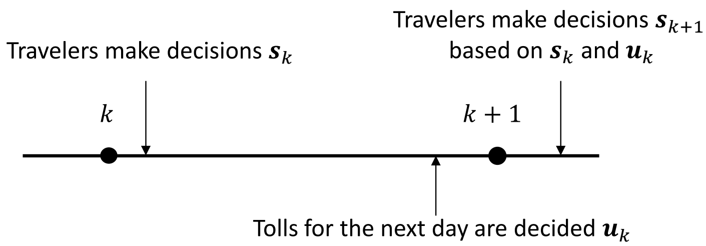

FIGURE 3: Timeline for the MDP

In this model, the transition probabilities depend on the current state, the tolls, and $\boldsymbol{\epsilon_k}$ that reflects the extent of irrationality of travelers. Each player is assumed to compute his/her disutility along a path using the revealed tolls and estimated travel time (which in reality may be obtained from some type of web/mobile application that stores past network conditions) assuming that the rest of the users continue on their paths chosen on the previous day. The Bellman's equation of optimality can thus be written as follows: $J_k(i) = \max_{\mathbf{u}_k \in \mathbf{U}} \sum_j p_{ij}^k(\mathbf{u}_k) J_{k+1}(j)$. The optimal values can be obtained using backward induction and since the transition probabilities are assumed to have closed-form expressions, the objective may be minimized over the set $\mathbf{U}$. The value of K should be carefully chosen so that the system would continue to remain at an equilibrium once the prices are revoked.

Consider the example in Figure 1. The boundary condition for state 4 is set to 1. Suppose that travelers use a logit learning model and on day $k+1$ they choose between the two alternatives with probabilities that depend on tolls for day $k+1$ and the travel times observed on day k, while assuming that other travelers remain on the paths chosen on day k. Let $\mathbf{u}_k = (u_k^T, u_k^B)$ represent the vector of tolls on the top and bottom links between nodes 1 and 2 in the network. On day 0, we assume that the network is in one of the four states with equal probability. Given below are some expressions for the transition probabilities (expressions for the other transition probabilities can be written in a similar way).

$$p_{11}^k(\mathbf{u}_k, \epsilon_k)) = \text{Probability that both travelers stay on the top path} \tag{19}$$

$$= \left(\frac{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right)^2 \tag{20}$$

$$p_{12}^k(\mathbf{u}_k, \epsilon_k)) = \text{Probability that both travelers move to the bottom path} \tag{21}$$

$$= \left(\frac{\exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right)^2 \tag{22}$$

$$p_{13}^k(\mathbf{u}_k, \epsilon_k)) = \text{Probability that only traveler 2 switches to the bottom path} \tag{23}$$

$$= \left(\frac{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right) \left(\frac{\exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right) \tag{24}$$

$$p_{14}^k(\mathbf{u}_k, \epsilon_k)) = \text{Probability that only traveler 1 switches to the bottom path} \tag{25}$$

$$= \left(\frac{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right) \left(\frac{\exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)}{\exp\left(\frac{-4-u_k^t}{\epsilon_k}\right) + \exp\left(\frac{-1-u_k^b}{\epsilon_k}\right)} \right) \tag{26}$$

In this example, we use the ϵ's defined in the previous case and solve the MDP to obtain the optimal pricing policy. The value of K was chosen to be 30 and the results are shown in Table 2. The transition probabilities (calculated based on the assumed ϵ's) imply that travelers in state 4 would continue to remain in that state with probability 1 (to within floating point accuracy) after 30 days. The following table shows the value functions. As can be seen from the first row of the table, the probability of reaching a UE at the end of the 30^{th} day is close to 1 from all states. Since on day 0, the system is equally likely to be in each of the four states, the long run percentage of reaching equilibrium is (0.994487 + 0.994637 + 0.993277 + 0.998216)/4 = 0.99515, which is a lot higher than long run percentages found in the previous cases. Note that the J's of the states for which the boundary values are set to 0 gradually increases as k decreases, while the J values for the UE state (for which the boundary value is initialized at 1) monotonically decreases as k decreases.

TABLE 2: Results of the MDP for $K = 30$

k	$J_k^*(1)$	$\boldsymbol{u}_k(1)$	$J_k^*(2)$	$\boldsymbol{u}_k(2)$	$J_k^*(3)$	$\boldsymbol{u}_k(3)$	$J_k^*(4)$	$\boldsymbol{u}_k(4)$
0	0.994487	(0,2)	0.994637	(1.3,0.1)	0.993277	(2,0)	0.998216	(0.1,1)
1	0.992994	(0,2)	0.993633	(1.2,0)	0.992195	(2,0)	0.999651	(0,1)
2	0.99122	(0,2)	0.992382	(1.1,0)	0.990809	(2,0)	0.999942	(0,1)
3	0.989275	(0,2)	0.990894	(1.1,0)	0.989083	(2,0)	0.99999	(0,1)
4	0.987079	(0,2)	0.989121	(1.8,0.7)	0.986993	(2,0)	0.999998	(0,1)
5	0.984541	(0,2)	0.98701	(1.5,0.4)	0.984503	(2,0)	1	(0,1)
6	0.981574	(0,2)	0.98451	(1.4,0.3)	0.981558	(2,0)	1	(0,1)
7	0.978133	(0,2)	0.981561	(1,0)	0.978125	(2,0)	1	(0,1)
8	0.97406	(0,2)	0.978127	(1,0)	0.974056	(2,0)	1	(0,1)
9	0.969226	(0,2)	0.974057	(1,0)	0.969224	(2,0)	1	(0,1)
10	0.963502	(0,2)	0.969225	(1,0)	0.963501	(2,0)	1	(0,1)
11	0.956699	(0,2)	0.963501	(1,0)	0.956698	(2,0)	1	(0,1)
12	0.948654	(0,2)	0.956698	(1,0)	0.948654	(2,0)	1	(0,1)
13	0.93907	(0,2)	0.948654	(1.4,0.4)	0.93907	(2,0)	1	(0,1)
14	0.927772	(0,2)	0.93907	(1,0)	0.927772	(2,0)	1	(0,1)
15	0.914254	(0,2)	0.927772	(1.4,0.4)	0.914254	(2,0)	1	(0,1)
16	0.898417	(0,2)	0.914254	(1.9,0.9)	0.898417	(2,0)	1	(0,1)
17	0.879299	(0,2)	0.898417	(1.9,0.9)	0.879299	(2,0)	1	(0,1)
18	0.857185	(0,2)	0.879299	(1.4,0.4)	0.857185	(2,0)	1	(0,0.9)
19	0.830007	(0,2)	0.857185	(1,0)	0.830007	(2,0)	1	(0,0.8)
20	0.799366	(0,2)	0.830007	(1.4,0.4)	0.799366	(2,0)	1	(0,0.7)
21	0.76033	(0,2)	0.799366	(1.4,0.4)	0.76033	(2,0)	1	(0,0.7)
22	0.718567	(0,2)	0.76033	(1.9,0.9)	0.718567	(2,0)	1	(0,0.6)
23	0.661377	(0,2)	0.718567	(1,0)	0.661377	(2,0)	1	(0,0.6)
24	0.606445	(0,2)	0.661377	(1.9,0.9)	0.606445	(2,0)	1	(0,0.5)
25	0.519531	(0,2)	0.606445	(1.9,0.9)	0.519531	(2,0)	1	(0,0.4)
26	0.453125	(0,2)	0.519531	(1.4,0.4)	0.453125	(2,0)	1	(0,0.4)
27	0.3125	(0,2)	0.453125	(1.4,0.4)	0.3125	(2,0)	1	(0,0.4)
28	0.25	(0,2)	0.3125	(1.4,0.4)	0.25	(2,0)	1	(0,0.3)
29	9.36E-14	(0,2)	0.25	(1,0)	7.86E-53	(0.1,0.2)	1	(0,0.3)
30	0	(0,0)	0	(0,0)	0	(0,0)	1	(0,0)

Table 3 reports the probability of reaching an UE state at the end of the pricing time periods for different time period durations, i.e., for different values of K. As expected, the longer we price the network, the chances of reaching an equilibrium increase.

TABLE 3: Results of the MDP for different values of K

K	$J_0^*(1)$	$J_0^*(2)$	$J_0^*(3)$	$J_0^*(4)$
5	0.605062	0.617078	0.517995	0.87233
10	0.83596	0.840394	0.799958	0.946907
15	0.92915	0.931078	0.913597	0.97707
20	0.969615	0.970441	0.962945	0.990166
25	0.987056	0.987408	0.984215	0.995811
30	0.994487	0.994637	0.993277	0.998216

Note that in the above example, we set the boundary value of state 4 to 1. Instead, we could have chosen state 3 as the preferred UE state and set its boundary value to 1. Although, states 3 and 4 have the same link flows, note that they are different path flow solutions. Hence, the proposed pricing mechanism not only improves the probability of reaching a UE solution but also lets us choose a particular UE solution in the presence of multiple equilibria. A preferred UE solution in such instances may be identified on the basis or stability or entropy. Further, one could set the boundary value of disequilibrium states (such as states 1 and 2) to 1. However, after the prices are revoked, since these states are not in equilibrium, travelers would continue to switch routes. In such cases, travelers may be forced to choose the disequilibrium state by enforcing a static toll that is collected indefinitely. This procedure may be used to help users reach a SO state.

5 CONCLUSION

In this paper, a dynamic pricing model was developed to improve the probability of convergence to a UE in the TAP with boundedly rational users. The problem was formulated as an MDP and was demonstrated using a small network and the numerical results appear promising. The problem presents several interesting directions for future research. Theoretical results related to rates of convergence of closed-form transition probabilities in the presence of dynamic pricing may be of interest. The dynamic pricing models may be extended to ones in which the transition probabilities are not known but are inferred from observed players' choices. These extensions can capture the effects of heterogeneity in learning and the value of time of travelers. It would also be interesting to see how these models work in practice by conducting experiments using simulations and human subjects.

6 ACKNOWLEDGMENT

This research was supported by the Data-Supported Transportation Operations and Planning University Transportation Center. The authors are grateful for this support.

REFERENCES

[1] J. Wardrop, “Some theoretical aspects of road traffic research,” *Proceedings of the Institution of Civil Engineers, Part II*, vol. 1, pp. 352–362, 1952.

[2] J. Nash, “Non-cooperative games,” *Annals of Mathematics*, vol. 54, no. 2, pp. pp. 286–295, 1951.

[3] A. C. Pigou, *The Economics of Welfare*. London: Macmillan and Co., 1920.

[4] H. Bar-Gera, “Origin-based algorithm for the traffic assignment problem,” *Transportation Science*, vol. 36, no. 4, pp. 398–417, 2002.

[5] R. B. Dial, “A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration,” *Transportation Research Part B*, vol. 40, no. 10, pp. 917–936, 2006.

[6] M. Frank and P. Wolfe, “An algorithm for quadratic programming,” *Naval Research Logistics Quarterly*, vol. 3, pp. 95–110, 1956.

[7] Y. Sheffi, *Urban Transportation Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[8] M. Mitradjieva and P. O. Lindberg, “The stiff is moving — conjugate direction Frank-Wolfe methods with application to traffic assignment,” *Transportation Science*, vol. 47, no. 2, pp. 280–293, 2013.

[9] R. Jayakrishnan, W. T. Tsai, J. N. Prashker, and S. Rajadhyaksha, “A faster path-based algorithm for traffic assignment,” *Transportation Research Record*, vol. 1443, pp. 75–83, 1994.

[10] T. Larsson and M. Patriksson, “Simplicial decomposition with disaggregated representation for the traffic assignment problem,” *Transportation Science*, vol. 26, pp. 4–17, 1992.

[11] E. Cascetta, “A stochastic process approach to the analysis of temporal dynamics in transportation networks,” *Transportation Research Part B: Methodological*, vol. 23, no. 1, pp. 1 – 17, 1989.

[12] E. Cascetta and G. E. Cantarella, “A day-to-day and within-day dynamic stochastic assignment model,” *Transportation Research Part A: General*, vol. 25, no. 5, pp. 277 – 291, 1991.

[13] G. E. Cantarella and E. Cascetta, “Dynamic processes and equilibrium in transportation networks: towards a unifying theory,” *Transportation Science*, vol. 29, no. 4, pp. 305–329, 1995.

[14] T. L. Friesz, D. Bernstein, N. J. Mehta, R. L. Tobin, and S. Ganjalizadeh, “Day-to-day dynamic network disequilibria and idealized traveler information systems,” *Operations Research*, vol. 42, no. 6, pp. 1120–1136, 1994.

[15] G. W. Brown, “Iterative solution of games by fictitious play,” in *Activity Analysis of Production and Allocation* (T. Koopmans, ed.), pp. 374–376, Wiley, New York, 1951.

[16] J. Robinson, “An iterative method of solving a game,” *Annals of Mathematics*, vol. 54, no. 2, pp. pp. 296–301, 1951.

[17] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical biosciences*, vol. 40, no. 1, pp. 145–156, 1978.

[18] J. M. Smith and G. Price, "The logic of animal conflict," *Nature*, vol. 246, p. 15, 1973.

[19] A. Nagurney and D. Zhang, "Projected dynamical systems in the formulation, stability analysis, and computation of fixed-demand traffic network equilibria," *Transportation Science*, vol. 31, no. 2, pp. pp. 147–158, 1997.

[20] G. W. Brown and J. Von Neumann, "Solutions of games by differential equations," tech. rep., DTIC Document, 1950.

[21] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT Press, 2010.

[22] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124 – 143, 1996.

[23] H. P. Young, "The evolution of conventions," *Econometrica*, vol. 61, no. 1, pp. pp. 57–84, 1993.

[24] H. P. Young, *Strategic learning and its limits*, vol. 2002. Oxford University Press, 2004.

[25] M. Kandori, G. J. Mailath, and R. Rob, "Learning, mutation, and long run equilibria in games," *Econometrica*, vol. 61, no. 1, pp. pp. 29–56, 1993.

[26] M. Kandori and R. Rob, "Evolution of equilibria in the long run: A general theory and applications," *Journal of Economic Theory*, vol. 65, no. 2, pp. 383–414, 1995.

[27] L. Blume, "Population games," Game Theory and Information 9607001, EconWPA, July 1996.

[28] M. J. Smith, "The existence, uniqueness and stability of traffic equilibria," *Transportation Research Part B: Methodological*, vol. 13, no. 4, pp. 295–304, 1979.