An optimal control approach to within day congestion pricing for stochastic transportation networks

Hemant Gehlot*[1], Harsha Honnappa†[2], and Satish V. Ukkusuri‡[1]

[1]Lyles School of Civil Engineering, Purdue University
[2]School of Industrial Engineering, Purdue University

Abstract

Congestion pricing has become an effective instrument for traffic demand management on road networks. This paper proposes an optimal control approach for congestion pricing for within day timescale that incorporates demand uncertainty and elasticity. Real time availability of the traffic conditions' information to the travelers and traffic managers allows periodic update of road pricing. We formulate the problem as an infinite-horizon countable-state Markov decision process (MDP) and analyze the problem to see if it satisfies conditions for conducting a satisfactory solution analysis. Such an analysis of MDPs is often dependent on the type of state space as well as on the boundedness of travel cost functions. We do not constrain the travel cost functions to be bounded and present an analysis centered around weighted sup-norm contractions that also holds for unbounded cost functions. We find that the formulated MDP satisfies a set of assumptions to ensure Bellman's optimality condition. Through this result, the existence of an optimal stationary policy for the MDP is shown. An approximation scheme is developed to resolve the implementation and computational issues of solving the control problem. Numerical results suggest that the approximation scheme efficiently solves the problem and produces accurate solutions.

Keywords— congestion pricing, within day timescale, demand elasticity, Bellman's optimality, countable states, optimal control

1 Introduction

Traffic demand management through congestion pricing (CP) has become one of the most important tools to mitigate congestion. The aim of CP is to change travelers' behavior by modifying the associated travel costs. In some cases, the aim is to modify the departure time by shifting some people from peak hours to off-peak times. In some cases, the goal is to discourage the use of particular routes by too many people and decrease the overall negative externalities related to congestion. The first CP was implemented in Singapore in 1975 and thereafter, it became popular in many parts of the world (Seik, 2000). There have been several works on CP in the literature. Some of these studies have focused on pricing models with an underlying static traffic assignment theory (De Palma and Lindsey, 2011) whereas others have modeled dynamic congestion pricing (Tan et al., 2015). Static models are simplified models that do not account for time-dependent behavior of traffic flows but are computationally less burdensome than dynamic models (Cheng et al., 2017). There is another way of classifying transportation models based on the timescale to which they are applied: day-to-day and within-day timescale models. Thus, the applications of CP can be also categorized into the aforementioned classes.

*Email: hgehlot@purdue.edu
†Email: honnappa@purdue.edu
‡Corresponding author. Email: sukkusur@purdue.edu. Address: Lyles School of Civil Engineering, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051.

In day-to-day modeling, CP is applied over a set of consecutive days but is only focused to a particular time period of the day whereas in within day models, CP is applied at various time periods of a day. Sandholm (2002) proposed a continuous CP for day-to-day route choice adjustment to guarantee efficient utilization of the system. Tan et al. (2015) incorporated user heterogeneity in day-to-day timescale CP. Rambha and Boyles (2016) studied CP in discrete time setting and considered stochastic day-to-day route choice behaviors. In within day timescale, Vickrey (1969) first proposed a CP model for a single bottleneck and subsequently many works followed. Value of time (VOT) converts toll charge into time units and is frequently used in CP models. Small and Yan (2001) considered a simple network of two links and studied CP for groups of people with different VOTs. Viti et al. (2003) developed a bilevel model of CP where the upper level problem is the place where tolls are set and aims at minimizing the system cost, and the lower level problem calculates the dynamic user-optimal response of the transport system (in terms of route choice and departure-time choice). Ramadurai et al. (2010) formulate the dynamic equilibrium conditions for a single bottleneck model with heterogeneous commuters as a linear complementarity problem. A lot of works have also focused on formulating either a bilevel programming or mathematical programming with equilibrium constraints models.

Most of the aforementioned works in CP assume that the demand at a particular time of a day is deterministic and inelastic. But demand variation from time to time, a critical feature of demand uncertainty, can significantly change the traffic condition and bring inefficiency to the network (Pi and Qian, 2017). In addition, these works do not capitalize on the recent findings from the analysis of large scale sensing and networked data about mobility patterns at the city scale. Recent findings (Hasan et al., 2013a,b; Yan et al., 2014; Zhong et al., 2016) find regularity of travel patterns in urban cities using cell phone, smart card data and social media data. Recent studies also find the variance of travel time patterns within a narrow range thereby allowing the better characterization of uncertainty. In recent times, there has been a steep rise in the availability of big data sources that complement traditional data collection technologies like cameras and loop detectors. For example, sources like GPS and cellular data can generate frequent mobility information (Song et al., 2010). A key big-data insight from this type of data is the repeated observations over a period of time that allows for a richer characterization of uncertainty and the mobility-demand processes. Some of the studies in real-time routing utilize the recent advancements in the technologies (Gao and Chabini, 2006; Pi and Qian, 2017). Most of these studies (e.g. Gao and Chabini (2006)) model optimal routing from the perspective of users rather than optimizing traffic demand as in pricing models. Pi and Qian (2017) considered a stochastic optimal approach from the system optimal perspective by minimizing the expected travel time of all travelers. However, Pi and Qian model an optimal routing problem from system optimal perspective without considering the selfish nature of the users. The former approach is efficient from the perspective of the system but the latter approach is fair from the perspective of users (Zhu and Ukkusuri, 2017). Thus, modeling of CP from the perspectives of both users and system is necessary to reach stable and efficient states in the network.

As mentioned before, the literature of CP models can be divided into static and dynamic modeling paradigms. Though dynamic models represent time variant flows, most of them are computationally burdensome (Ukkusuri et al., 2012). Even though the academic research in traffic assignment has moved towards dynamic modeling, static models remain widely popular for strategic transport planning due to their efficiency and simplicity (Bliemer et al., 2012). Thus, we formulate a *pseudodynamic* modeling approach wherein traffic flows across the network in an epoch or time step are considered to be static, and once the demand in that epoch is realized, these flows can be determined by ensuring that travel times across all routes of an OD pair are equal. We assume that the traffic generated in each time step dissipates in that time step and there is no spillover to the next time step, which is similar to the assumption made for *quasi-dynamic network loading*, where it is assumed that traffic propagates instantaneously through the network in a single time step (Bliemer et al., 2012). The state of the system is characterized by the travel demand in each time step, which is modeled as a random variable to represent the associated uncertainty. The dependencies between consecutive time steps come through demand elasticity. The parameters of the demand distribution are governed by the traffic conditions (total system travel time) in the last time step. Since traffic conditions in a time step is a function of the demand generated in that step plus the pricing actions taken, this problem can be formulated as a Markov decision process (MDP).

We formulate the problem as a countable state MDP, whose details are described in the next section. Bellman's equation, which represents the optimality condition for MDPs is the basic entity in MDP theory and almost all existence, characterization and computational results are based on analysis of it. Unlike the finite space case, the satisfaction of Bellman's condition by countable state MDPs requires satisfaction of additional conditions. These assumptions ensure that the required contraction properties of the Bellman operator are satisfied. If these assumptions are not satisfied then one or more of the following may happen: the optimality equations may not have a solution, an optimal policy might not exist, an optimal policy may exist but it might not be stationary (time-invariant), and standard algorithms may fail to converge to the optimal solution (Bertsekas, 2012a). Apart from the type of state space, the sets of assumptions also depend on whether the cost per time step is bounded or unbounded, which in turn is dependent on whether the underlying travel cost functions are bounded or unbounded. The solution analysis simplifies if the cost functions are assumed to be bounded and/or the state space is finite[1]. For example, if we consider MDPs with countable state space and bounded costs then it is sufficient to show that the Bellman operator is a contraction with respect to the *unweighted* sup-norm. However, travel cost functions proposed by the US Bureau of Public Roads are not bounded (Boyce et al., 1981), and consequently we do not assume bounded cost functions and present an analysis using *weighted sup-norm contractions* that also holds for unbounded costs (Bertsekas, 2012b). To the best knowledge of authors, this is the first such analysis in optimal control modeling of CP.

A natural question that arises is whether the formulated MDP satisfies the assumptions to ensure existence of an optimal policy? In addition, while implementing a solution algorithm how should we compute and store the optimal policy for an infinite number of states? In the literature, approximation schemes are proposed to implement countable state MDPs, so we also develop one such approach (Aviv and Federgruen, 1999). Also, the solution computation of MDPs becomes burdensome with the size of the problem. In such problems, ordinary linear algebra operations such as n-dimensional inner products, are prohibitively consuming, and indeed it may be even impossible to store a n-vector in a computer memory. Hence, it is necessary devise an efficient solution algorithm for solving the MDP. Once we efficiently find the solutions of MDP, the question becomes what is the optimal policy/rule that the traffic managers should follow? Should large tolls be put during low demand periods? This paper makes a significant contribution to gain insights into the these research questions.

In summary, this study makes the following contributions:

- A countable-state average cost MDP of congestion pricing is formulated for within day timescale that incorporates demand stochasticity and elasticity.
- The developed formulation does not constrain the travel cost functions to be bounded, leading to an elaborate solution analysis using the theory of weighted sup-norm contractions and recurrence properties of Markov chains. We analytically show that the MDP satisfies sufficient conditions to ensure Bellman's optimality conditions, implying the existence of an optimal stationary policy.
- Develop an approximation scheme to efficiently compute the solutions of the problem.

This study is organized in the following manner. The next section describes the details of the problem. Section 3 presents the solution methodology used in this paper. Section 4 presents the conducted numerical results. The final section concludes the study and provides future directions.

2 Problem Statement

Consider a traffic network with a single origin-destination (OD) pair (we extend the analysis to multiple OD networks in Appendix B). It is assumed that traffic demand in each time step is a random variable but the probability distribution of the demand is a function of the traffic conditions of the previous

[1]For instance, Rambha and Boyles (2016) considered day-to-day CP modeling (for a particular time of the day) to find the optimal tolling policy for the average cost problem that distributes travelers across different routes for a fixed and deterministic demand. The state space is finite in their problem, and the solution analysis is simpler since an analysis centered around contractions is not required.

time step and the corresponding tolling actions, reflecting demand elasticity. The information of traffic conditions is available to the people through technologies like advanced traveler information systems (ATIS) and more recent advancements in big data sources like GPS, cellular data etc. Static traffic flow conditions are modeled within every time step, so the demand generated in each time step is resolved in that time step (Bliemer et al., 2012). Note that we will use the terms *congestion pricing* and *tolling* interchangeably in this paper. The tolling agency periodically decides to take tolling actions on various paths across the OD pair in a day. At the start of each time step, the agency decides how much toll to assign across different routes based on the traffic flow conditions. The information of traffic flow conditions is regularly collected through technologies like cameras, loop detectors etc. It is assumed that these instruments cover a significant portion of the road network to provide the details of traffic conditions with good accuracy. The objective of the tolling agency is to minimize the average total system travel time over all the time steps during a day. This problem can be formulated as a controlled Markov chain, also known as Markov decision process (MDP) (Puterman, 2014). We model an infinite horizon problem which is a reasonable approximation for problems involving a finite but a very large number of time steps (Bertsekas, 2012a). For instance, Srinivasan et al. (2006) approximated the infinite horizon problem by conducting a 24-hour simulation of the traffic signal control problem. In addition, infinite horizon problems are interesting because their analysis is elegant and insightful, and the implementation of optimal policies is often simple. For example, optimal policies are typically stationary, i.e., the optimal rule for choosing control does not change from one time step to the next (Bertsekas, 2012a). Such a time-invariant optimal rule can also be easily applied by traffic managers and practitioners in comparison to time-specific rules. Due to the aforementioned reasons, the assumption of infinite horizon is frequently made in control literature such as traffic signal control (De Oliveira and Camponogara, 2010; Diakaki et al., 2002; Srinivasan et al., 2006).

2.1 State space

State at time step k is the total traffic demand x_k across the OD pair during that time step. It is assumed that the demand is Poisson distributed (Clark and Watling, 2005; Cova and Johnson, 2002; Hellerstein, 1991). So, the state space S is countable and takes integer values in the interval $[0, \infty)$.

2.2 Action space

Actions are the amount of toll to be levied across the routes in the network. The action vector at time step k is represented as $u_k \equiv \{u_k^1, \ldots, u_k^R\}$, where R is the number of routes between the OD pair. The elements of the action vector take finite values from the set $(\tau_1, \ldots, \tau_m)$. The minimum and maximum toll values are denoted by τ_{min} and τ_{max}, respectively, such that $\tau_{min} > 0$ and $\tau_{max} < \infty$. The action space is assumed to be state dependent and denoted as $U(x_k)$, at time step k.

2.3 Transition probabilities

The travel demand in each time step is assumed to be a random variable since the total demand can be stochastic during different times of a day (Pi and Qian, 2017). To ensure demand elasticity, we model the Poisson process with a state-dependent rate equal to $\frac{\theta}{TSTT(x_k, u_k)}$. It follows that the likelihood of observing large demand in the next time step decreases if the total system travel time ($TSTT$) of the current time step increases and vice versa. Denote $p_{x_k x_{k+1}}(u_k)$ as the probability to go to state x_{k+1} in the time step $k+1$ if the state at time step k was x_k and action u_k was applied. The transition probability $p_{x_k x_{k+1}}(u_k)$ is given by

$$p_{x_k x_{k+1}}(u_k) = \frac{e^{-\frac{\theta}{TSTT(x_k, u_k)}} \left(\frac{\theta}{TSTT(x_k, u_k)}\right)^{x_{k+1}}}{x_{k+1}!}, \tag{1}$$

where θ is a given positive constant.

2.4 Costs

The cost function $g(x_k, u_k)$ is the expected cost per time step when decision u_k is taken in state x_k and is given as follows,

$$g(x_k, u_k) = \sum_{x_{k+1} \in S} p_{x_k x_{k+1}}(u_k) TSTT(x_{k+1}, u_k).$$

2.5 Objective

We define a stationary policy μ as a policy/rule of choosing actions that does not change from one time step to the next. A stationary policy μ is a *feasible policy* if $\mu(x) \in U(x)$, $\forall x \in S$. Also, $J_\mu(x)$ is defined as the average cost per time step or the expected $TSTT$ for policy μ assuming that the system starts at state x, i.e. $x_0 = x$,

$$J_\mu(x) = \lim_{K \to \infty} \frac{1}{K} E\left\{ \sum_{k=0}^{K-1} g(x_k, \mu(x_k)) | x_0 = x \right\}$$

where $E\{\cdot\}$ is the expected value operator. The objective is to find an optimal policy μ^* such that

$$J^*(x) \equiv J_{\mu^*}(x) = \min_{\mu \in \Pi} J_\mu(x),$$

where $\prod$ is the space of all feasible policies.

We now present the formulation of the total system travel time ($TSTT$) in the MDP.

2.6 Total system travel time

Recall that traffic conditions within a time step are assumed to be static. We consider an affine cost function $c^r(x) = c^r x$ for route r and toll u_k^r corresponding to route r at time step k. We start with the assumption that there is no correlation/overlap between different routes; however, we extend the analysis to networks with overlapping routes in Appendix A. In the literature, unbounded polynomial cost functions as adopted by the US Bureau of Public Roads of the form $c^r(x) = c^r x^a + b$, where a and b are positive constants, are considered (Boyce et al., 1981). Assuming the cost functions to be affine may nevertheless be a good approximation for most nonlinear models (Altman, 1994; Correa et al., 2008; Roughgarden and Tardos, 2002; Steinberg and Zangwill, 1983). Steinberg and Zangwill (1983) presented the necessary and sufficient conditions for Braess' Paradox to occur under affine cost assumptions. Roughgarden and Tardos (2002) analytically characterized the degradation of network performance due to unregulated traffic in terms of price of anarchy for networks with linear cost functions. Walters (1961) reports that, for flows along roads below capacity, empirical evidence supports such an approximation.

We assume that the total demand x_k of a time step gets distributed across different routes such that travel time across all the routes is the same, in accordance with Wardrop's principle (while distributing the flow across different routes we assume that flow is infinitesimally divisible to simplify the expression of total system travel time). This ensures that, while satisfying the objective of minimizing the total system travel time, the traffic flows also represent users' perspective. The solution of the following equations[2] provides the expression for $TSTT$

$$\frac{TSTT(x_k, u_k)}{R} = c^1 x_k^1 + u_k^1 = \ldots = c^R x_k^R + u_k^R$$

and

$$x_k^1 + \ldots + x_k^R = x_k,$$

where x_k^r denotes traffic flow across route r at time step k.

[2]The formulated model can accommodate non-linear costs but will require numerically solving a set of non-linear equations to compute the expression of $TSTT$ (Liu, 1974).

Therefore, the expression for $TSTT$ is

$$TSTT(x_k, u_k) = Rx_k \frac{\prod_{\forall r \in R} c^r}{\sum_{\forall r \in R} \prod_{\forall j \neq r} c^j} + R \frac{\sum_{\forall r \in R} \left(\left(\prod_{\forall j \neq r} c^j \right) u_k^r \right)}{\sum_{\forall r \in R} \prod_{\forall j \neq r} c^j}. \tag{2}$$

3 Solution methodology

In general, the analysis of average cost problems with denumerable states and unbounded cost per time step poses many difficulties, and at present there is no comprehensive theory. However, there are sets of assumptions that allow a satisfactory analysis (Bertsekas, 2012a). We first present these assumptions and later show that the MDP satisfies them.

3.1 Bellman's optimality condition

We assume that state 0 is special in that the system has a tendency to return to it under all policies. Such a state comes under the category of *recurrent* states. In particular, for any policy μ we denote C_μ as the expected cost starting from state 0 up to the first return to 0 and N_μ as the expected number of time steps to return to state 0 starting from 0 and we assume the following.

Assumption 1. *For every policy μ, C_μ and N_μ are finite. Furthermore, N_μ is uniformly bounded over μ, i.e., for some $\overline{N} > 0$, we have $N_\mu < \overline{N}$ for all μ.*

We now introduce a positive sequence $v = \{v_0, v_1, \ldots\}$, such that

$$\inf_{i=0,1,\ldots} v_i > 0,$$

and the weighted sup-norm

$$||J|| = \max_{i=0,1,\ldots} \frac{|J(i)|}{v_i}$$

in the space $B(S)$ of sequences $\{J(0), J(1), \ldots\}$ such that $||J|| < \infty$. As stated before, weighted sup-norm contractions play an important role in the solution analysis of the MDPs that have countable state space and unbounded cost per time step (Bertsekas, 2012b). The following assumptions form an essential part of that analysis.

Assumption 2. *The sequence $G = \{G_0, G_1, \ldots\}$, where*

$$G_i = \max_{u \in U(i)} |g(i, u)|, \quad i = 0, 1, \ldots$$

belongs to $B(S)$.

Assumption 3. *The sequence $V = \{V_0, V_1, \ldots\} \in B(S)$, where*

$$V_i = \max_{u \in U(i)} \sum_{j=0}^{\infty} p_{ij}(u) v_j, \quad i = 0, 1, \ldots$$

belongs to $B(S)$.

Assumption 4. *There is a scalar $\rho \in (0, 1)$ and an integer $m \geq 1$ such that for all μ and $i = 0, 1, \ldots,$ we have*

$$\frac{\sum_{j=1}^{\infty} p_{ij}^m(\mu) v_j}{v_i} \leq \rho \tag{3}$$

where $p_{ij}^m(\mu)$ is the probability of entering state j at the mth time step given that the state at time step 0 was i and policy μ is applied.

The following proposition provides the main result.

Proposition 1. Let Assumptions 1, 2, 3 and 4 hold. Then the optimal average cost, denoted λ^*, is the same for all initial states and together with some sequence $h^* = \{h^*(0), h^*(1), \ldots\}$ satisfies Bellman's equation

$$h^*(i) + \lambda^* = \min_{u \in U(i)} \left[g(i,u) + \sum_{j=0}^{\infty} p_{ij}(u) h^*(j) \right], \quad i = 0, 1, \ldots \tag{4}$$

Furthermore, if $\mu^*(i)$ attains the minimum in the above equation for all i, the stationary policy μ^* is optimal.

Proof. See Proposition 5.6.5 in volume 2 of Bertsekas (2012a). □

We now show that our MDP satisfies the aforementioned assumptions under some conditions. First, we introduce the definition of a positive recurrent state (Brémaud, 2013).

Definition 1. A state in a Markov chain is positive recurrent if the expected number of time steps to return to the same state is finite.

We now introduce some results that will be useful in checking if the MDP satisfies Assumption 1.

Lemma 1. (Foster's theorem) Consider an irreducible discrete-time Markov chain on a countable state space S having a transition probability matrix with elements p_{ij} for going from state i to state j, where $i, j \in S$. Then, a Markov chain is positive recurrent if and only if there exists a Lyapunov function $v : S \to \mathbb{R}$, such that $v_i \geq 0, \forall i \in S$ and

$$\sum_{j \in S} p_{ij} v_j < \infty \quad \forall i \in F \tag{5}$$

$$\sum_{j \in S} p_{ij} v_j < v_i - \epsilon \quad \forall i \notin F \tag{6}$$

for some finite set F and strictly positive ϵ.

Proof. See Theorem 1.1, Chapter 5 of Brémaud (2013). □

We now present an expression that will be useful for proving Assumption 1 and other results in the paper.

Lemma 2. $g(x_k, \mu_k) = \frac{k^0 \lambda}{TSTT(x_k, \mu_k)} + \sum_r k^r \mu_k^r$

Proof. Recall that

$$g(x_k, \mu_k) = \sum_{x_{k+1}=0}^{\infty} p_{x_k x_{k+1}}(\mu_k) TSTT(x_{k+1}, \mu_k), \tag{7}$$

where $TSTT(x_k, \mu_k) = k^0 x_k + \sum_r k^r \mu_k^r$, such that k^0 and k^r are positive constants coming from the Equation (2). Note that we denote $\mu_k(x_k) \equiv \mu_k$ to simplify the notation. Expanding the expression for TSTT,

$$g(x_k, \mu_k) = \sum_{x_{k+1}=0}^{\infty} p_{x_k x_{k+1}}(\mu_k) \left(k^0 x_{k+1} + \sum_r k^r \mu_k^r \right) \tag{8}$$

$$= k^0 E\{x_{k+1}\} + \left(\sum_{x_{k+1}=0}^{\infty} p_{x_k x_{k+1}}(\mu_k) \right) \left(\sum_r k^r \mu_k^r \right) \tag{9}$$

$$= k^0 \frac{\lambda}{TSTT(x_k, \mu_k)} + \sum_r k^r \mu_k^r \tag{10}$$

as $E\{x_{k+1}\} = \frac{\lambda}{TSTT(x_k, \mu_k)}$ and $\sum_{x_{k+1}=0}^{\infty} p_{x_k x_{k+1}}(\mu_k) = 1$. □

We now start showing that the MDP satisfies Assumption 1. We first show that expected recurrence times are finite for all feasible policies.

Lemma 3. *N_μ is finite in the MDP for all the feasible policies.*

Proof. The Foster's theorem requires the MDP to be irreducible, which is the case as all the pairs of states i and j are communicating since $p_{ij}(\mu) > 0, \forall \mu$. Denote $v_i = TSTT(i, u^{max})$, where $u^{max} \equiv \{\tau^{max}, \dots, \tau^{max}\}$. So, $TSTT(i, u^{max})$ is the total system travel time for state i when the toll values on all the routes are maximum and equal to τ^{max}. First, we show that condition (5) is satisfied for all the states and feasible policies of the MDP. For state i and policy μ,

$$\begin{aligned}\sum_{j\in S} p_{ij}(\mu)v_j =& \sum_{j\in S} p_{ij}(\mu) TSTT(j, u^{max}) \\ =& \sum_{j\in S} p_{ij}(\mu)(k^0 j + \sum_r k^r \tau^{max}) \\ =& \frac{k^0\theta}{TSTT(i,\mu)} + \sum_r k^r \tau^{max}\end{aligned}$$

The last equality in the above set of equations comes after doing an analysis similar to the proof of Lemma 2. Let us compute the maximum value of $\sum_{j\in S} p_{ij}(\mu)v_j$ over all the states and policies and show that it is finite. Note that

$$\max_\mu \left(\max_i \left(\sum_{j\in S} p_{ij}(\mu)v_j\right)\right) = \frac{k^0\theta}{TSTT(0, u^{min})} + \sum_r k^r\tau^{max} < \infty \tag{11}$$

where $TSTT(0, u^{min})$ is the total system travel time for state 0 when the toll values on all the routes are minimum and equal to τ^{min}. Note that the inequality in Equation (11) holds true as $\tau_{min} > 0$.

We now show that there exists a finite set of states such that all the elements not belonging to this set satisfy the condition (6). We claim that for all the states $i \ge \sqrt{\frac{\theta}{k^0}} + 1$, condition (6) is satisfied with $\epsilon = k^0 \left(\sqrt{\frac{\theta}{k^0}} + 1 - \frac{\theta}{k^0\left(\sqrt{\frac{\theta}{k^0}}+1\right)}\right)$. Here, F is set of all integers in the interval $[0, \sqrt{\frac{\theta}{k^0}} + 1)$. For a policy μ, we compute

$$\begin{aligned}& v_i - \sum_{j\in S} p_{ij}(\mu)v_j \\ &= k^0 i + \sum_r k^r\tau^{max} - \sum_{j\in S} p_{ij}(\mu)\left(k^0 j + \sum_r k^r \tau^{max}\right) \\ &= k^0 i + \sum_r k^r\tau^{max} - \frac{k^0\theta}{k^0 i + \sum_r k^r\mu^r} - \sum_r k^r\tau^{max} \\ &= k^0 i - \frac{k^0\theta}{k^0 i + \sum_r k^r\mu^r} \\ &> k^0 i - \frac{k^0\theta}{k^0 i}\end{aligned}$$

Note that the function $k^0 i - \frac{k^0\theta}{k^0 i}$ becomes equal to zero at $i = \sqrt{\frac{\theta}{k^0}}$ and is monotonically increasing with i as the first-order derivative of the function is $k^0 + \frac{\theta}{i^2} > 0, \forall i$. So, for all policies and for states $i \ge \sqrt{\frac{\theta}{k^0}} + 1$,

$$v_i - \sum_{j\in S} p_{ij}(\mu)v_j > k^0\left(\sqrt{\frac{\theta}{k^0}} + 1 - \frac{\theta}{k^0\left(\sqrt{\frac{\theta}{k^0}}+1\right)}\right)$$

Thus, from Lemma 1, $N_\mu < \infty, \forall \mu$. □

Before showing that the expected recurrence times N_μ are uniformly bounded we state the following basic limit theorem of Markov chains.

Lemma 4. *Consider a recurrent irreducible aperiodic Markov chain. Let p_{ii}^k be the probability of entering state i at the kth time step given that the initial state (state at time step 0) is i. Also, by convention $p_{ii}^0 = 1$. Let f_{ii}^k be the probability of first returning to state i at the kth time step. Thus,*

$$p_{ii}^k - \sum_{l=0}^{k} f_{ii}^{k-l} p_{ii}^l = \begin{cases} 1, & \text{if } k = 0 \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\lim_{k\to\infty} p_{ii}^k = \frac{1}{\sum_{k=0}^{\infty} k f_{ii}^k}.$$

Proof. See Theorem 1.2, chapter 3 of Karlin (2014). □

Denote $f_{ii}^k(\mu)$ as the probability of first returning to state i at the kth time step given that the initial state was i and when policy μ was applied. We first show that $f_{00}^k(\mu)$ is non-zero for all feasible policies μ of the MDP, which will be useful in showing that the expected recurrence times are uniformly bounded.

Lemma 5. *$f_{00}^k(\mu)$ is non-zero for all feasible policies μ of the MDP.*

Proof. The probability of first returning to state 0 after the first time step given that the initial state was 0 and policy μ was applied is given by

$$f_{00}^1(\mu) = P(x_1 = 0|x_0 = 0, \mu),$$

where $P(x_{k+1} = j|x_k = i, \mu)$ is the probability of entering state j at time step $k+1$ given that the state at time step k is i and policy μ is applied and is equal to $p_{ij}(\mu)$. Similarly, we compute the first return probability at time step 2,

$$f_{00}^2(\mu) = P(x_2 = 0, x_1 \neq 0|x_0 = 0, \mu) \tag{12}$$
$$= P(x_2 = 0|x_1 \neq 0, \mu)P(x_1 \neq 0|x_0 = 0, \mu). \tag{13}$$

The equality in Equation (13) comes because of the Markov property (Brémaud, 2013). Following in the same way, the first return probability after k time steps can be written as

$$f_{00}^k(\mu) = P(x_k = 0|x_{k-1} \neq 0, \mu) \prod_{m=2}^{k-1} P(x_m \neq 0|x_{m-1} \neq 0, \mu)P(x_1 \neq 0|x_0 = 0, \mu).$$

Note that

$$\begin{aligned} P(x_k = 0|x_{k-1} \neq 0, \mu) &= \frac{P(x_k = 0, x_{k-1} \neq 0, \mu)}{P(x_{k-1} \neq 0, \mu)} \\ &= \frac{\sum_{i>0} P(x_k = 0|x_{k-1} = i, \mu)P(x_{k-1} = i|\mu)}{\sum_{j>0} P(x_{k-1} = j|\mu)} \\ &= \sum_{i>0} P(x_k = 0|x_{k-1} = i, \mu)\overline{P}(x_{k-1} = i|\mu), \end{aligned}$$

where $\overline{P}(x_{k-1} = i|\mu) = \frac{P(x_{k-1}=i|\mu)}{\sum_{j>0} P(x_{k-1}=j|\mu)}$ and $P(x_k = 0|x_{k-1} = i, \mu) = e^{-\frac{\theta}{TSTT(i,\mu)}}$. Similarly,

$$\begin{aligned} P(x_k \neq 0|x_{k-1} \neq 0, \mu) &= 1 - P(x_k = 0|x_{k-1} \neq 0, \mu) \\ &= 1 - \sum_{i>0} P(x_k = 0|x_{k-1} = i, \mu)\overline{P}(x_{k-1} = i|\mu). \end{aligned}$$

Also,

$$P(x_1 \neq 0|x_0 = 0, \mu) = 1 - P(x_1 = 0|x_0 = 0, \mu)$$
$$= 1 - e^{-\frac{\theta}{TSTT(0,\mu)}}.$$

So, it follows that

$$f_{00}^k(\mu) = \left(\sum_{i>0} e^{-\frac{\theta}{TSTT(i,\mu)}} \overline{P}(x_{k-1} = i|\mu)\right).$$
$$\prod_{m=2}^{k-1}\left(1 - \sum_{i>0} e^{-\frac{\theta}{TSTT(i,\mu)}} \overline{P}(x_{m-1} = i|\mu)\right)\left(1 - e^{-\frac{\theta}{TSTT(0,\mu)}}\right).$$

We now show that $f_{00}^k(\mu)$ is positive. The following holds

$$1 - e^{-\frac{\theta}{TSTT(0,\mu)}} > 0, \forall i, \mu,$$

because $\frac{\theta}{TSTT(0,\mu)} > 0, \mu$. Also,

$$\sum_{i>0} e^{-\frac{\theta}{TSTT(i,\mu)}} \overline{P}(x_{m-1} = i|\mu) < \sum_{i>0} \overline{P}(x_{m-1} = i|\mu),$$

as $e^{-\frac{\theta}{TSTT(i,\mu)}} < 1, \forall i, \mu$. Note that

$$\sum_{i>0} \overline{P}(x_{m-1} = i|\mu) = \sum_{i>0} \frac{P(x_{m-1} = i|\mu)}{\left(\sum_{j>0} P(x_{m-1} = j|\mu)\right)} = 1.$$

Therefore,

$$\prod_{m=2}^{k-1}\left(1 - \sum_{i>0} e^{-\frac{\theta}{TSTT(i,\mu)}} \overline{P}(x_{m-1} = i|\mu)\right) > 0.$$

Also, since we have shown that $P(x_k = 0|x_{k-1} \neq 0, \mu) > 0$, so $P(x_{k-1} \neq 0, \mu) > 0$ for $P(x_k = 0|x_{k-1} \neq 0, \mu)$ to remain defined. So, $\exists i \neq 0$, such that $\overline{P}(x_{k-1} = i|\mu) > 0$. Thus, combining the fact that $e^{-x} > 0 \ \ \forall x$, we get

$$\sum_{i>0} e^{-\frac{\theta}{TSTT(i,\mu)}} \overline{P}(x_{k-1} = i|\mu) > 0.$$

Therefore, $f_{00}^k(\mu) > 0, \forall \mu$. □

We now show that the expected recurrence times of the MDP are uniformly bounded.

Lemma 6. *For our MDP, there exists a $\overline{N} > 0$, such that $N_\mu < \overline{N}, \forall \mu$.*

Proof. We prove this lemma through the method of contradiction. Let us assume that the lemma is not true. Then, there exists a sequence of policies $\{\mu_j\}$ with $j = 0, 1, \dots$ such that for this sequence the mean recurrence times $\{N_{\mu_j}\}$ form a non-decreasing sequence such that

$$\lim_{j\to\infty} N_{\mu_j} = \infty.$$

Note that

$$\lim_{j\to\infty} \sum_{l=0}^{k} f_{00}^{k-l}(\mu_j) p_{00}^l(\mu_j) = \sum_{l=0}^{k} f_{00}^{k-l}(\mu_\infty) p_{00}^l(\mu_\infty),$$

and the sum is well-defined since the limits are probabilities. For a policy μ_j, $N_{\mu_j} = \sum_{k=0}^{\infty} k f_{00}^k(\mu_j)$. Using the fact that $\lim_{j\to\infty} N_{\mu_j} = \infty$, it follows from Lemma 4 that

$$\lim_{k\to\infty} \sum_{l=0}^{k} f_{00}^{k-l}(\mu_\infty) p_{00}^l(\mu_\infty) = 0.$$

But, this can only happen if every term in the sum is zero (observe that each term is non-negative). Therefore, it follows that

$$f_{00}^{k-l}(\mu_\infty)p_{00}^{l}(\mu_\infty) = 0 \quad \forall k \geq 0, l \leq k.$$

Since $p_{00}^0(\cdot) = 1$ by definition as in Lemma 4, it must be the case that $f_{00}^k = 0$. But this is a contradiction of Lemma 5. Thus, our assumption that the expected recurrence times are not uniformly bounded is false, completing the proof. □

Lemma 7. *For our MDP, C_μ is finite for all policies μ.*

Proof. Denote C'_μ as the expected cost starting from state 0 up to the first return to 0 when the number of time steps of return is K. Then,

$$C'_\mu = \sum_{k=0}^{K-1} g\left(x_k, \mu_k\right)$$

such that $x_0 = 0, x_{K-1} = 0$. From Lemma 2 we get,

$$\begin{aligned} C'_\mu &= \sum_{k=0}^{K-1} \left(k^0 \left(\frac{\theta}{k^0 x_k + \sum_r k^r \mu_k^r} \right) + \sum_r k^r \mu_k^r \right) \\ &< \sum_{k=0}^{K-1} \left(k^0 \left(\frac{\theta}{\sum_r k^r \mu_k^r} \right) + \sum_r k^r \mu_k^r \right). \end{aligned}$$

Note that $k^0 \left(\frac{\theta}{\sum_r k^r \mu_k^r} \right) + \sum_r k^r \mu_k^r < \infty$ as all the possible toll values are positive and bounded (i.e. $\tau^{min} > 0$ and $\tau_{max} < \infty$). Denote φ^{max} as the upper bound of $k^0 \left(\frac{\theta}{\sum_r k^r \mu_k^r} \right) + \sum_r k^r \mu_k^r$ for all policies μ. Now we take expectation of C'_μ over the number of time steps K to obtain C_μ. So, for policy μ

$$\begin{aligned} C_\mu <& E_K \left\{ \sum_{k=0}^{K-1} \left(k^0 \left(\frac{\theta}{\sum_r k^r \mu_k^r} \right) + \sum_r k^r \mu_k^r \right) \right\} \\ \leq& E_K \{ K \varphi^{max} \} \\ =& N_\mu \varphi^{max} \\ \leq& \overline{N} \varphi^{max} < \infty. \end{aligned}$$

Thus, C_μ is finite for all the policies. □

Proposition 2. *Assumption 1 is satisfied in the MDP.*

Proof. Lemmas 3, 6 and 7 imply satisfaction of Assumption 1. □

We now show that the MDP satisfies Assumption 2.

Proposition 3. *There exists a sequence $v = \{v_0, v_1, \ldots\}$ that satisfies Assumption 2 for all the instances of the MDP.*

Proof. The expression for G_i is given as

$$G_i = \max_{u \in U(i)} |g(i,u)| = TSTT(i, u^{max})$$

Now we choose $v_i = TSTT(i, u^{max})$. Notice that $\inf_{i=0,1,\ldots} TSTT(i, u^{max}) = TSTT(0, u^{max}) > 0$ as all the coefficients in Equation (2) are positive and maximum toll value τ^{max} is also positive. Then,

$$\frac{|G_i|}{v_i} = 1 \quad \forall i$$

Therefore, $||J|| < \infty$. □

Next, we show that Assumption 3 is satisfied by MDP.

Proposition 4. *There exists a sequence $v = \{v_0, v_1, \ldots\}$ that satisfies Assumption 3 for all the instances of the MDP.*

Proof. Set $v_i = TSTT\left(i, u^{max}\right)$. Then,

$$\begin{aligned} V_i &= \max_{u \in U(i)} \sum_{j=0}^{\infty} p_{ij}(u) v_j \\ &= \max_{u \in U(i)} \sum_{j=0}^{\infty} p_{ij}(u) TSTT(j, u^{max}) \\ &= \max_{u \in U(i)} \sum_{j=0}^{\infty} p_{ij}(u) \left(k^0 j + \sum_r k^r \tau^{max} \right) \end{aligned}$$

Since $\sum_{j=0}^{\infty} p_{ij}(u) j = \frac{\theta}{TSTT(i,u)}$ and $\sum_{j=0}^{\infty} p_{ij}(u) = 1$, it follows that

$$V_i = \max_{u \in U(i)} \left(\frac{k^0 \theta}{TSTT(i,u)} + \sum_r k^r \tau^{max} \right) = \frac{k^0 \theta}{TSTT(i, u^{min})} + \sum_r k^r \tau^{max},$$

where $TSTT(i, u^{min})$ is the total system travel time for state i when the toll values on all the routes are minimum and equal to τ^{min}. Then,

$$E_i = \frac{|V_i|}{v_i} = \frac{\frac{k^0 \theta}{TSTT(i,u^{min})} + \sum_r k^r \tau^{max}}{TSTT(i, u^{max})}, \tag{14}$$

as $V_i > 0 \; \forall i$ holds for all the parameters of the MDP. In order to find $\max_{i=0,1,\ldots} \frac{V_i}{v_i}$, we differentiate the expression in Equation (14) with respect to i to get

$$\frac{\partial E_i}{\partial i} = \frac{-k^0 \theta \left(k^0 TSTT(i, u^{min}) + k^0 TSTT(i, u^{max}) \right)}{\left(TSTT(i, u^{min}) TSTT(i, u^{max}) \right)^2} - \frac{k^0 \sum_r k^r \tau^{max}}{(TSTT(i, u^{max}))^2} < 0,$$

for all the states i and parameter values. Thus, E_i is a decreasing function. Hence, the maximum value of the expression E_i occurs at $i = 0$. That is,

$$\max_{i=0,1,\ldots} \frac{V_i}{v_i} = \frac{\frac{k^0 \theta}{TSTT(0,u^{min})} + \sum_r k^r \tau^{max}}{TSTT(0, u^{max})} < \infty,$$

as all the parameter values are finite and the minimum toll value $\tau^{min} > 0$. □

Finally, we have the following result:

Proposition 5. *If* $\frac{k^0 \theta}{\sum_r k^r \tau^{min}} \leq e^{-\left(\frac{\theta}{\sum_r k^r \tau^{min}} \right)} \sum_r k^r \tau^{max}$ *then the MDP satisfies Assumption 4.*

Proof. Let $m = 1$ in condition (3). Then, left hand-side (LHS) of condition (3) becomes

$$\frac{\sum_{j=1}^{\infty} p_{ij}(\mu) v_j}{v_i}.$$

Using the relation $\sum_{j=1}^{\infty} p_{ij}(\mu) v_j = \sum_{j=0}^{\infty} p_{ij}(\mu) v_j - p_{i0}(\mu) v_0$, we get

$$\frac{\sum_{j=1}^{\infty} p_{ij}(\mu) v_j}{v_i} = \frac{\sum_{j=0}^{\infty} p_{ij}(\mu) v_j - p_{i0}(\mu) v_0}{k^0 i + \sum_r k^r \tau^{max}} = \frac{\sum_{j=0}^{\infty} p_{ij}(\mu) TSTT(j, u^{max}) - p_{i0}(\mu) TSTT(0, u^{max})}{k^0 i + \sum_r k^r \tau^{max}}.$$

Since $TSTT(j,u^{max}) = k^0 j + \sum_r k^r \tau^{max}$, $\sum_{j=0}^{\infty} p_{ij}(\mu) j = \frac{\theta}{TSTT(i,\mu)}$ and $\sum_{j=0}^{\infty} p_{ij}(\mu) = 1$, we get

$$\frac{\sum_{j=0}^{\infty} p_{ij}(\mu) TSTT(j,u^{max}) - p_{i0}(\mu) TSTT(0,u^{max})}{k^0 i + \sum_r k^r \tau^{max}}$$
$$= \frac{\frac{k^0\theta}{TSTT(i,\mu)} + \sum_r k^r \tau^{max}}{k^0 i + \sum_r k^r \tau^{max}} - \frac{p_{i0}(\mu)\left(\sum_r k^r \tau^{max}\right)}{k^0 i + \sum_r k^r \tau^{max}}. \quad (15)$$

There are two set of variables in the above expression: i and μ. We first find the minimum value of the expression with respect to both the variables through partial differentiation. Denote $\chi = \frac{k^0\theta}{TSTT(i,\mu)} + \sum_r k^r \tau^{max}$, $\kappa = -p_{i0}(\mu)(\sum_r k^r \tau^{max})$ and $\zeta = k^0 i + \sum_r k^r \tau^{max}$. We first compute the partial derivatives of these terms with respect to i. Note that

$$\frac{\partial \chi}{\partial i} = \frac{-(k^0)^2\theta}{(TSTT(i,\mu))^2} < 0.$$

Then,

$$\begin{aligned}
\frac{\partial \kappa}{\partial i} &= -\left(\sum_r k^r \tau^{max}\right) \frac{\partial p_{i0}(\mu)}{\partial i} \\
&= -\left(\sum_r k^r \tau^{max}\right) \frac{\partial e^{-\left(\frac{\theta}{TSTT(i,\mu)}\right)}}{\partial i} \\
&= -\left(\sum_r k^r \tau^{max}\right) e^{-\left(\frac{\theta}{TSTT(i,\mu)}\right)} \left(\frac{k^0\theta}{(TSTT(i,\mu))^2}\right) < 0.
\end{aligned}$$

Also,

$$\frac{\partial \zeta}{\partial i} = k^0 > 0.$$

Thus, maxima of χ and κ as well as the minimum of ζ all occur when $i = 0$, for all the policies and parameter settings. Therefore,

$$\max_{i=0,1,\dots} \frac{\sum_{j=1}^{\infty} p_{ij}(\mu) v_j}{v_i} = \frac{\frac{k^0\theta}{TSTT(0,\mu)} + \sum_r k^r \tau^{max} - p_{00}(\mu)\left(\sum_r k^r \tau^{max}\right)}{k^0 0 + \sum_r k^r \tau^{max}}$$

as $\zeta = k^0 i + \sum_r k^r \tau^{max}$ is always positive. We now compute the maximum of the expression in RHS of Equation (15) with respect to μ. Taking partial derivatives of the right-hand side (RHS) expression in Equation (15) with respect to μ^r $\forall r$, to get

$$-\frac{\frac{k^0 k^r \theta}{(TSTT(i,\mu))^2}}{k^0 i + \sum_r k^r \tau^{max}} - \frac{e^{-\frac{\theta}{TSTT(i,\mu)}} \left(\frac{\theta k^r}{(TSTT(i,\mu))^2}\right) \left(\sum_r k^r \tau^{max}\right)}{k^0 i + \sum_r k^r \tau^{max}} < 0.$$

That is, the expression of Equation (15) is a decreasing function with respect to $\mu^r, \forall r$. Therefore, the maximum of the expression occurs at $\mu = \{\tau^{min}, \tau^{min}, \dots, \tau^{min}\}$, for all states and parameter settings.

Thus, by setting $i = 0$ and $\mu = u^{min}$ we bound the expression in Equation (15) as follows:

$$\frac{\frac{k^0\theta}{TSTT(i,\mu)} + \sum_r k^r \tau^{max}}{k^0 i + \sum_r k^r \tau^{max}} - \frac{p_{i0}(\mu)\left(\sum_r k^r \tau^{max}\right)}{k^0 i + \sum_r^R k^r \tau^{max}} < 1 + \frac{\frac{k^0\theta}{TSTT(0,u^{min})} - p_{00}(u^{min})\left(\sum_r k^r \tau^{max}\right)}{\sum_r k^r \tau^{max}}.$$

Since the supremum of ρ in condition (3) is 1, we ensure that the expression in the last line is less than or equal to 1 (as $\rho \in (0,1)$). That is,

$$1 + \frac{\frac{k^0\theta}{TSTT(0,u^{min})} - p_{00}(u^{min})\left(\sum_r k^r \tau^{max}\right)}{\sum_r k^r \tau^{max}} \le 1$$

Or,

$$\frac{k^0\theta}{\sum_r k^r\tau^{min}} \leq e^{-\left(\frac{\theta}{\sum_r k^r\tau^{min}}\right)} \sum_r k^r\tau^{max}$$

□

Finally, it follows that Proposition 1 holds for the MDP.

3.2 Implementation issues in solving the problem

We now focus on implementing a solution algorithm to solve the MDP. A few studies have established the convergence of algorithms like value iteration and policy iteration for countable state MDPs (Aviv and Federgruen, 1999; Meyn, 1997; Sennott, 1991). But the countable state MDP, while being theoretically appealing, faces several implementation issues. First, the infinite sums to the right of Equation (4) cannot be evaluated on a term by term basis, nor can the function $h^*(i)$ be evaluated and stored for infinitely many values (Aviv and Federgruen, 1999). Therefore, approximation algorithms need to be devised to find the solutions of countable state MDPs.

One approach to approximate a countable state MDP is by solving the finite state problem obtained through truncation of the state space. Existing results show that as the size of the approximating MDP increases, its cost function and, under some conditions its optimal policies approach those of the original, countable MDP (Altman, 1994; Cavazos-Cadena, 1986). So, for solving the problem we truncate the state space S to a finite interval $[0, x^{max}]$. Once we solve the finite state MDP we can extrapolate the optimal policies for the states $S\backslash[0, x^{max}]$. The truncation of the state space implies the state transition probabilities are now governed by a truncated Poisson distribution. We denote the transition probability from state x_k to state x_{k+1} for action u_k using truncated Poisson distribution as $q_{x_k x_{k+1}}(u_k)$, where

$$q_{x_k x_{k+1}}(u_k) = \frac{\frac{e^{-\frac{\theta}{TSTT(x_k,u_k)}}\left(\frac{\theta}{TSTT(x_k,u_k)}\right)^{x_{k+1}}}{x_{k+1}!}}{\sum_{X=0}^{X=x^{max}} \frac{e^{-\frac{\theta}{TSTT(x_k,u_k)}}\left(\frac{\theta}{TSTT(x_k,u_k)}\right)^{X}}{X!}}.$$

The action space, costs and objective function remain the same as before. We now present the solution algorithm that we use to solve the truncated MDP.

3.3 Solution algorithm

We use policy iteration to solve for the optimal policy of the infinite horizon finite state MDP. Policy iteration algorithm starts with a stationary policy μ^0, and iteratively generates a sequence of new policies $\mu^1, \mu^2, \ldots$ as follows:

1. (Policy evaluation) Given a policy μ^m at iteration m, we compute average and differential costs λ^m and $h^m(i)$ satisfying

$$h^m(i) + \lambda^m = g(i, \mu^m(i)) + \sum_{j=0}^{x^{max}} q_{ij}(\mu^m(i))h^m(j), \quad i = 0, \ldots, x^{max}.$$

2. (Policy improvement) We find a stationary policy μ^{m+1}, where for all i, $\mu^{m+1}(i)$ is such that

$$g(i, \mu^{m+1}(i)) + \sum_{i=0}^{x^{max}} q_{ij}(\mu^{m+1}(i))h^m(j) = \min_{u \in U(i)} \left[g(i, u) + \sum_{i=0}^{x^{max}} q_{ij}(u)h^m(j) \right].$$

3. (Termination check) If $\lambda^{m+1} = \lambda^m$ and $h^{m+1}(i) = h^m(i)$ for all i, the algorithm terminates; otherwise, we go to Step 1 with μ^{m+1} replacing μ^m.

3.4 Aggregated problem

We have seen that dynamic programming (DP) algorithms like Policy iteration can be used to solve the finite state MDP. However, the computational requirements associated with the exact solution of even finite state MDPs are overwhelming. It is known that DP algorithms can be numerically applied only if the sizes of the state space are relatively small. Computing the exact solution is generally difficult and possibly intractable for large problems due to the widely known *curse of dimensionality* (Bertsekas, 2012a). Thus, it is necessary to devise an approximation scheme to tackle this issue.

We use state aggregation to overcome the large state space. Before performing the aggregation, we assume that the state space is continuous by using the normal approximation to the Poisson distribution. So, the probability density function of going to the state x_{k+1} at time step $k+1$ if the state in the previous time step was x_k and the action u_k was taken is given by

$$f_{x_k x_{k+1}}(u_k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_{k+1}-\upsilon)^2}{2\sigma^2}},$$

where $\upsilon = \frac{\theta}{TSTT(x_k,u_k)}, \sigma^2 = \frac{\theta}{TSTT(x_k,u_k)}$.

3.4.1 State space

We construct a finite number of states from the continuous state space by aggregating states using intervals. We define the aggregated state space set $S' = \left\{ \left[0, \frac{x^{max}}{N}\right], \left[\frac{x^{max}}{N}, \frac{2x^{max}}{N}\right], \ldots, \left[\frac{(N-1)x^{max}}{N}, x^{max}\right] \right\}$ by dividing x^{max} into N intervals. For an aggregated interval/state $y \in S'$ we approximate the values in the interval by its center, y^c.

3.4.2 Transition probabilities

We compute the state transition probability of going from an aggregated state X to another aggregated state Y by computing the probability of going from the center of state X to state Y. Since we truncate the state space in the interval $[0, x^{max}]$, it may happen that the sum of probabilities of going from state X to all the aggregated states might not add up to one. Thus, we define

$$p'_{y_k^c y_{k+1}^c}(u_k) = \int_{y_{k+1}^c - \frac{x^{max}}{2N}}^{y_{k+1}^c + \frac{x^{max}}{2N}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\upsilon)^2}{2\sigma^2}} dx,$$

where $\upsilon = \frac{\theta}{TSTT(y_k^c,u_k)}, \sigma^2 = \frac{\theta}{TSTT(y_k^c,u_k)}$. We now normalize these values by setting

$$p_{y_k^c y_{k+1}^c}(u_k) = \frac{p'_{y_k^c y_{k+1}^c}(u_k)}{\sum_{\forall y^c \in S'} p'_{y_k^c y^c}(u_k)}.$$

Note that the action space for the approximate dynamic programming remains the same as before. The expressions for costs also remain the same as before and only involves replacing the states with centers of the aggregated states. Thus, the policy iteration algorithm can be used similarly as in Section 3.3 by appropriately using the state space and transition probabilities of the aggregated problem.

4 Numerical results

Consider a single origin-destination (OD) pair with two routes as shown in Figure 1. In the appendices, we show how networks with complex topologies can be reduced to simple single OD pair networks with non-overlapping routes. So, we present the results for the simplest setting. The cost parameters c^i for the routes are 1 and 2 units for the routes 1 and 2, respectively. The action space is a vector consisting of three toll values: 2, 3 and 4 units. The parameter θ of the Poisson distribution is set to 100. The maximum value of demand per unit time x_{max} is set to 20. We term this problem as the *original* problem in this paper. The results that we present will either use the *original* problem or modifications of it.

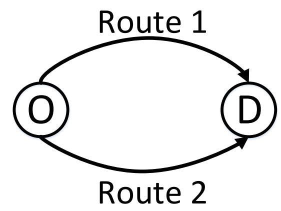

Figure 1: Test network

4.1 Solutions

We now present the solutions obtained from solving the *original* problem. We first solve the problem as described in Section 3.3. Figure 2 provides the optimal toll values for the two routes. The dashed line presents the optimal toll values for route 1 and the solid line provides the toll values for route 2. It can be seen that for both the routes the optimal value of toll is larger for small state values as compared to larger state values. This is possibly because when the demand is low, putting large toll value reduces the probability of a large demand in the next time step by Equation (1). This switching of optimal policy is also dependent on the parameters θ and $k^0, \ldots, k^r$. To better understand this, we analyze the expected cost per time step $g(x_k, \mu_k)$. We partially differentiate the expression $g(x_k, \mu_k)$ in Lemma 2 with respect to $\mu_k^r, \forall r$. On differentiating the expression with different routes r, we get the same expression for different routes and find that the derivative of $g(x_k, \mu_k)$ with respect to μ_k^r is equal to zero when $k^0 x_k + \sum_{r \in R} k^r \mu_k^r = \sqrt{k^0 \theta}$ (we also checked the second order conditions to verify that the first order conditions represent the minimum). Consider a network of two routes with route cost parameters equal to c^1 and c^2, for simplicity. Then, the derivative of $g(x_k, \mu_k)$ with respect to μ_k^r is equal to zero when

$$\sqrt{\frac{c^2}{c^1}}\mu_k^1 + \sqrt{\frac{c^1}{c^2}}\mu_k^2 = \sqrt{\frac{\theta}{c^1 c^2}} - x_k. \tag{16}$$

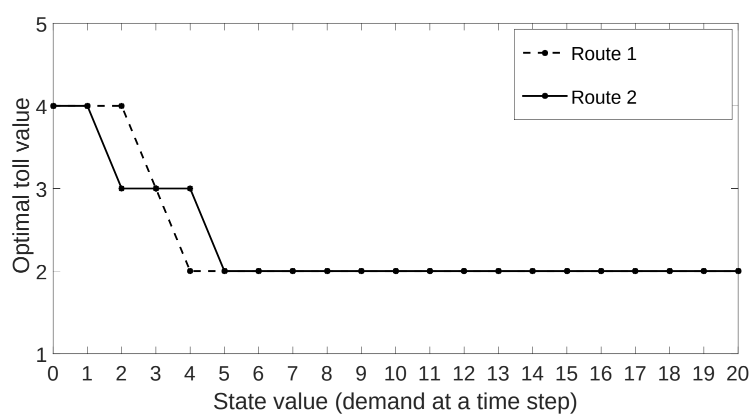

Figure 2: Plot of optimal toll values for different states by solving the *original* problem.

First, we consider the case when $\sqrt{\frac{\theta}{c^1 c^2}} - x_k$ is positive. We plot the condition (16) in Figure 3. If the current demand is low (i.e., $x_k < \sqrt{\frac{\theta}{c^1 c^2}}$) then the optimal toll values satisfy Equation (16) as the LHS of equation is always positive. For a given set of the parameters, optimal toll value increases as the demand decreases by Equation (16). In the extreme case, where the length of perpendicular from the origin to the line denoting Equation (16) becomes very large, i.e. $\frac{c^1 c^2}{\sqrt{(c^1)^4 + (c^2)^4}} \left(\sqrt{\frac{\theta}{c^1 c^2}} - x_k \right) >$

$\sqrt{(\tau^{max})^2 + (\tau^{max})^2}$, then it is optimal to put the largest possible toll value τ_{max} so that the difference between the LHS and RHS of the Equation (16) is minimum. We now consider the case when $\sqrt{\frac{\theta}{c^1 c^2}} - x_k$ is not positive as shown in Figure 4. Since the cost parameters and the possible toll values are positive, the LHS cannot become negative and thus it is optimal to put the least toll value τ^{min} so that the difference between the LHS and RHS is minimum. Thus, the optimal toll values are governed by the interplay between the demand x_k and parameters θ, c^1 and c^2.

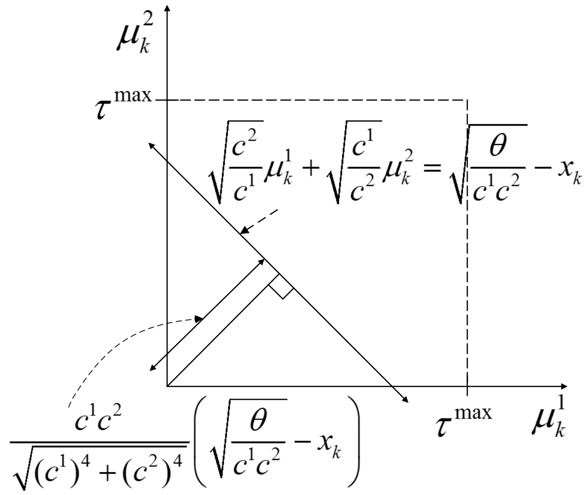

Figure 3: Plot of the first order condition of the expected cost per unit time when $\sqrt{\frac{\theta}{c^1 c^2}}$ is positive.

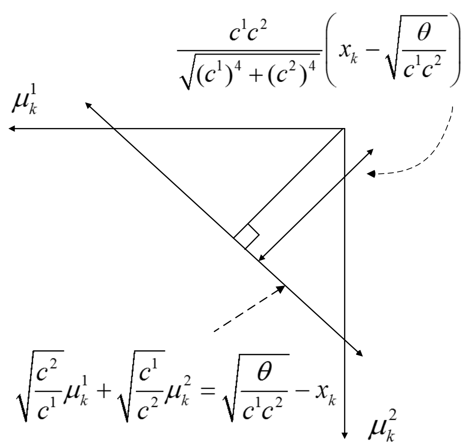

Figure 4: Plot of the first order condition of the expected cost per unit time when $\sqrt{\frac{\theta}{c^1 c^2}}$ is not positive.

We now present some numerical results to illustrate the point in the last paragraph. Figure 5 presents the variation of optimal toll values with θ, by keeping the cost parameters to be fixed. It can be seen that as θ increases the optimal toll value for a given demand value also increases. This can also be followed from the analysis of expected cost $g(x_k, \mu_k)$: as θ increases, the toll values μ_k^1 and μ_k^2 also increase by Equation (16), for a given demand value and cost parameters. This is also what one would anticipate: when θ is low, the toll should be low, since the demand in the next time step is going to be small with large probability by Equation (1); when θ is large, the demand in the next time step is likely to be large, and therefore charging a large toll is appropriate. Thus, optimal toll calculation is anticipatory in nature as it takes into account the possible demand distribution of the next time step.

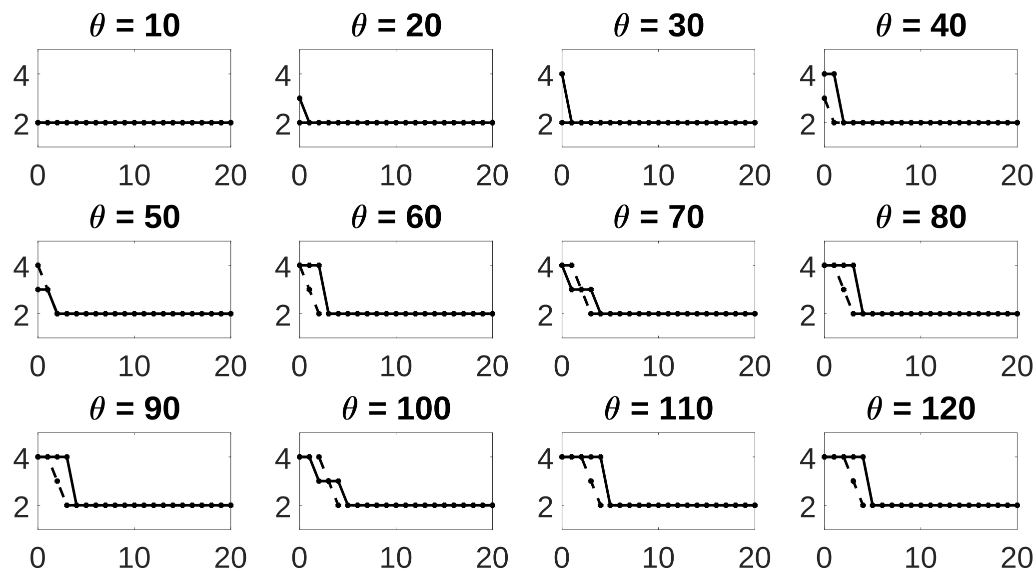

Figure 5: Plot of optimal toll values by solving the *original* problem for different values of θ. The x-axis represents state value (demand value) at a particular time step and y-axis represents the optimal toll value for that time step. The dashed and solid lines present optimal toll values for routes 1 and 2, respectively.

4.2 Truncation

In Section 3.2, we introduced truncation of the states as a strategy to solve the infinite state problem. We claimed that as the truncation value (x^{max}) becomes large the solution of the truncated problem can be approximated to the solution of the untruncated problem. In Figure 6, we plot the optimal average cost λ^* obtained from solving the *original* problem but for different values of x^{max}. It can be seen that λ^* converges to a fixed value when x^{max} becomes equal to 20. Thus, choosing x^{max} equal to 20 is a reasonable assumption. At the same time, solving the MDP through truncation scheme always provides an approximate solution of the countable state MDP. For instance, we observed that for large values of the parameter θ, the optimal policy becomes to put low toll values, contrary to the argument provided in the previous section. But this effect disappears if we choose a larger truncation value x^{max}. At the same time, the choice of a larger truncation value x^{max} comes at the cost of a significantly increased computation time (shown in the next section). Thus, there is a trade-off between solution quality versus computation time when we make the choice of truncation value.

4.3 Variation of computation time with the number of states

In Section 3.4, we stated that as the number of states in the problem increase the computational time for solving the problem significantly increases. Figure 7 presents the computational time when the *original* problem is solved by varying the values of x^{max}. Note that the number of states in the problem are equal to $1+x^{max}$ because the state values start from zero. So, increasing the value of x^{max} implies increasing the number of states in the problem. The pattern of increase in the computational time with the number of states is non-linear from the figure. Therefore, truncated problem needs to

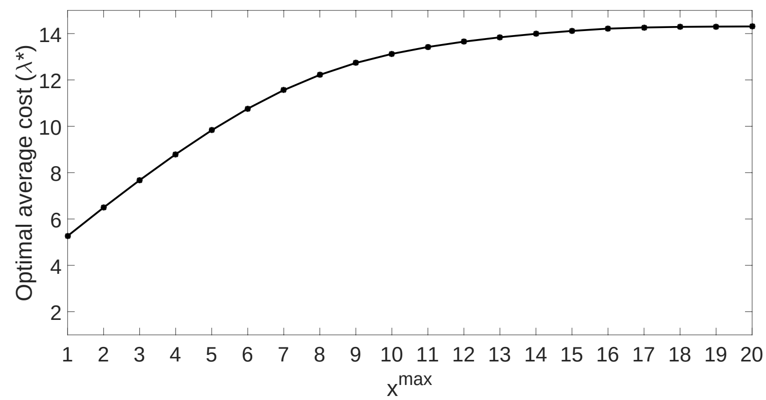

Figure 6: Variation of optimal average cost with the state truncation value (x^{max}).

be approximated so that the solutions of the problem can be efficiently solved. So, from this point we present the results by solving an approximation of the *original* problem.

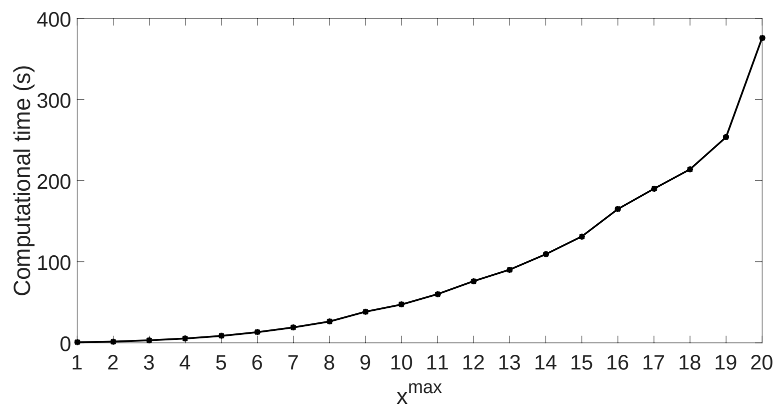

Figure 7: Variation of computational time (in seconds) with the state truncation value (x^{max}).

4.4 Variation of optimal average cost with the number of aggregated states

Before we perform aggregation of the problem as suggested in the last section, we compare the solution quality of the aggregated problem with the *original* problem. Figure 8 presents the comparison of the optimal average cost obtained from solving the *original* and aggregated problems. The points on the solid line in the figure present the value of optimal average cost for various number of aggregated states in the aggregated problem. The dashed line represents the optimal average cost obtained from solving the *original* problem. It can be seen that the solution of aggregated problem approaches to the solution of the *original* problem as the number of aggregated states increase. Based on these tests, we decided to solve the subsequent results with 10 aggregated states. The computational time associated with solving the problem with 10 aggregated states is about 12 seconds, which is about 18 times lower than solving the *original* problem. Thus, the aggregated problems can be efficiently solved.

4.5 Variation of computational time with the number of routes

In the *original* problem, we considered two routes connecting the OD pair. We analyze the variation of computational time for solving the problem as the number of routes increase. Figure 9 presents this

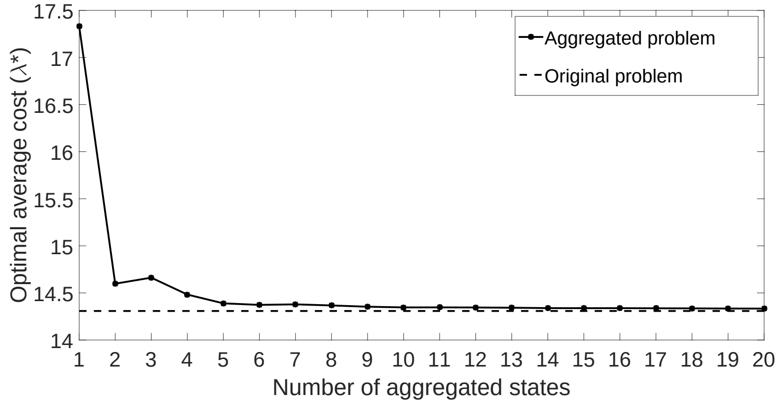

Figure 8: Variation of optimal average cost with the number of aggregated states.

variation when the aggregated problem is solved. It can be seen that computational time significantly increases with the number of routes. This is in accordance with network modeling problems that face increasing computational times with network size (Gehlot and Ukkusuri, 2018). A possible remedy to tackle this issue in future works can be to aggregate routes similarly as we aggregate states in the MDP.

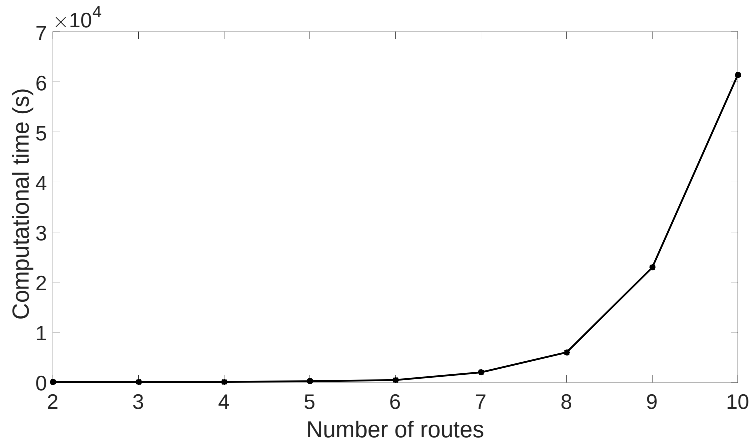

Figure 9: Variation of computational time (in seconds) with the number of routes

5 Conclusions and future directions

In this paper, we propose an optimal control modeling approach for within day timescale congestion pricing that incorporates demand elasticity and stochasticity. We develop a countable-state infinite-horizon MDP where we do not constrain the travel cost functions to be bounded, requiring a more elaborate solution analysis centered around weighted sup-norm contractions and recurrence properties of Markov chains. We prove that an optimal policy exists by verifying sufficient conditions needed to ensure Bellman's optimality. The developed model is both efficient from the system perspective and fair from users' perspective. We also develop an approximation scheme that resolves implementation and computational issues associated with solving the exact problem. Through our conducted numerical tests we find that the optimal tolling calculation is anticipatory of the demand distribution in the next time step and sometimes it is optimal to levy large tolls whereas sometimes it is optimal to levy low tolls, depending on the problem parameters. It is also found that the approximate algorithm scheme is both efficient and accurate.

There can be several extensions to this study. We did not include learning from travels on past days and incorporating it will also be a meaning full extension. Also, we assumed a stationary (time-invariant) model by keeping the parameter θ to be the same across different time steps. However, if we assume time-varying parameters then it is notably harder to analyze the non-stationary model; see Bowerman (1974) for the finite state space setting. Computing the optimal policy is complicated as well; in practice one may have to use rolling/planning horizon methods (Alden and Smith, 1992) or use classic linear programming methods (Ghate and Smith, 2013) when the state space is finite. Indeed, there is very little literature on how to deal with non-stationary MDPs with unbounded costs and countable state spaces - two critical features of our model. This promises to be a fruitful avenue for further research. Finally, considering stochasticity in the supply side will also be an interesting future study.

Acknowledgement

The authors are grateful to National Science Foundation for the grant CMMI 1520338 to support the research presented in this paper.

A Extension to networks with overlapping routes

In Section 2.6, we derived the expression of $TSTT$ for a single OD pair with independent/non-overlapping routes in terms of the various parameters. We now analyze the problem with overlapping routes. Notice that in order to conduct the analysis in Section 3, we used the the fact that total system travel time satisfies a particular form, i.e. $TSTT(x, u) = k^0 x + \sum_r k^r u_k$, with positive parameters k^0, k^r. We now analyze if the same form holds when there are overlaps in the routes of a network. We present an algorithm that can be used to compute the expression of $TSTT$ for single OD pair networks with overlapping routes. Note that a network is a composition of two types of connections: *series* and *parallel*. A set of elements (e.g. links) are connected in *series* if the same amount of traffic flows through them. Individual elements in the network are said to be in *parallel* if they are connected between the same pair of nodes. In Figure 10, consider a network with origin A, destination D and two overlapping routes connecting the OD pair: 1) link 1-link 2-link 4, and 2) link 3-link 4. In this network, the same traffic flows through links 1 and 2, so they are connected in series. Link 3 and the combined set of links 1 and 2 are in *parallel* as they are connected between the same pair of nodes, A and C. Finally, link 4 and the combined set of links 1, 2 and 3 are connected in *series*.

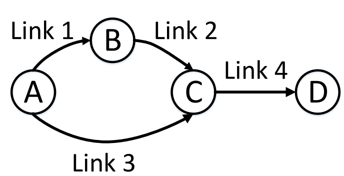

Figure 10: A network with overlapping routes

We now present some results that we will use in computing the expression for $TSTT$.

Lemma 8. *Consider a pair of nodes that are connected by a route with R consecutive links. The links have cost parameters as $c^1, \dots, c^R$ and toll values as $u^1, \dots, u^R$, respectively. Then, the total system travel time of this network is the same as that of a single link connecting the pair of nodes with cost parameter equal to $\sum_{r=1}^R c^r$ and toll equal to $\sum_{r=1}^R u^r$. Thus, R links of the route can be replaced by a single link of cost parameter $\sum_{r=1}^R c^r$ and toll equal to $\sum_{r=1}^R u^r$.*

Proof. Denote x as the travel demand flowing between the pair of nodes. Then, the travel times on links $1, \dots, R$ are $c^1 x + u^1, \dots, c^R x + u^R$, respectively. So, the total system travel time of the network is $(c^1x+u^1)+\ldots+(c^Rx+u^R)=(c^1+\ldots+c^R)x+(u^1+\ldots+u^R)$, which is also equal to travel time of a single link with cost parameter equal to $\sum_{r=1}^{R} c^r$ and toll equal to $\sum_{r=1}^{R} u^r$. $\square$

Lemma 9. *Consider a pair of nodes that have R non-overlapping routes connecting the OD pair. The routes have cost parameters as $c^1,\ldots,c^R$ and tolls as $u^1,\ldots,u^R$, respectively. Then, the total system travel time of this network is the same as that of a single link connecting the pair of nodes with cost equal to $R\frac{\prod\limits_{\forall r\in R} c^r}{\sum\limits_{\forall r\in R}\prod\limits_{\forall j\neq r} c^j}$ and toll equal to $R\frac{\sum\limits_{\forall r\in R}\left(\left(\prod\limits_{\forall j\neq r} c^j\right)u_k^r\right)}{\sum\limits_{\forall r\in R}\prod\limits_{\forall j\neq r} c^j}$.*

This proof of this result comes trivially from the expression (2), so we do not present here for brevity. We now present the algorithm to reduce a network with overlapping routes to a single link network that has the same $TSTT$ as the original network:

1. Denote U and V as the upstream and downstream nodes in the current iteration, respectively.
2. Initialize the upstream node as the origin of the network.
3. Pick an outgoing route from node U. Move along this route until a node is found that is an intersection of multiple routes. Denote this node as V.
4. Find all the elements that are connected to the nodes U and V in parallel. Using Lemmas 8 and 9, replace all the elements such that only single links are connected in parallel between the nodes U and V.
5. Replace all the parallel links joining U and V with a single link by Lemma 9.
6. Replace all the links (in series) connecting the origin and node V with a single link by Lemma 8.
7. Set U equal to V. If U is the destination node then stop. Otherwise, go to step 3.

The correctness of the above algorithm can be proved by a loop invariant scheme, with the loop invariant being *at a particular iteration of the algorithm, there is a single link with a linear cost function connecting the origin and U.*

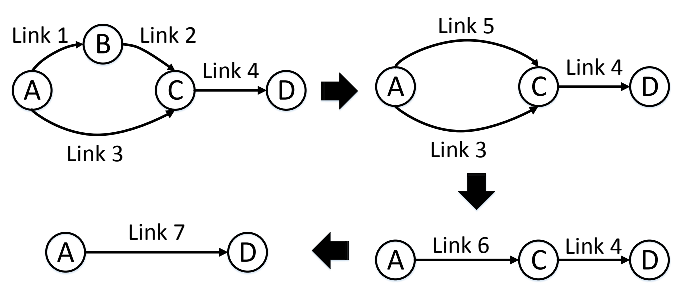

Figure 11: Reduction of an example network with overlapping routes to a single link network

To demonstrate how this algorithm works, consider Figure 11: 1) First, initialize U as node A. Then, node C is node V. Then, we replace links 1 and 2 with a single link, say link 5 using Lemma 8. Now links 5 and 3 are connected in parallel. So, links 5 and 3 can be replaced by a single link, say link 6. In the next iteration, nodes C and D become the U and V nodes, respectively. Finally, links 6 and 4 can be replaced by a single link (say link 7) using Lemma 8. So, a single link with cost of the form $k^0x+\sum_r k^r u^r$ connects the nodes A and D. Thus, we get the expression for total system travel time for this network. We can see that by using Lemmas 8 and 9, the cost parameters and the coefficients corresponding to toll values remain positive. Thus, the constants k^0 and k^r in the expression $TSTT = k^0x+\sum_r k^r u^r$ for networks with overlapping routes are positive. Thus, overlapping networks have $TSTT$ of the same form as non-overlapping networks and the analysis of sufficiency conditions that we earlier performed carries over.

B Extension to multiple OD pair networks

We now demonstrate that the MDP model carries over to multiple OD-pair networks. We consider a special case of multiple OD pairs, which we term as *series activity trips* networks. Consider the network in Figure 12 with three OD pairs: O-D_1,O-D_2 and O-D_3. At time step t, x_t is the number of people traveling from node O to different destinations: $\rho_1 x_t$ people travel to node D_1; of the remaining $(1-\rho_1)x_t$ people, a fraction ρ_2 of them travel to node D_2 and the remaining $(1-\rho_1-\rho_2+\rho_1\rho_2)\,x_t$ people travel to node D_3. Such type of networks are relevant for modeling interrelated choices underlying trip chaining behavior. Kitamura (1984) expresses the destination choice decision that underlies trip chaining as a series of sequential choices. Suppose a traveler is about to make a visit to another location after completing a visit at the current location. It may happen that some travelers might not travel to the other location because of an unexpected incidence (e.g. coming across a friend) or the purpose of travel is satisfied at the current location. But for some travelers the purpose may not be satisfied at the current location (e.g. some stock is unavailable at the first location in a shopping trip, so some people may go to another shopping location), so they will travel to the next destination. Some recent studies also indicate that series trip chaining behavior happens in electric vehicles (EVs) due to different electric-charing opportunities along a route (Tamor et al., 2013; Xie et al., 2017).

We characterize the state of the system by the total demand across different OD pairs. The demand in each time step changes by the transition probability function of Equation (1). Consider a

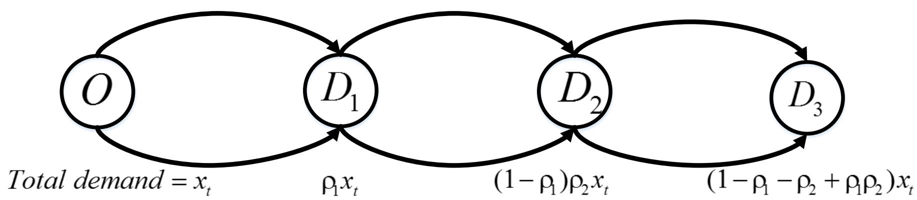

Figure 12: An example of a *series activity trips* network

network with total demand at time step t as x_t, then the demands across different OD pairs are some proportions of the total demand as shown in Figure 12. Under this setting, we show that the total system travel time at time step t is given by the same form as before, i.e. $TSTT = k^0 x_t + \sum_r k^r u_t^r$. We again reduce the whole network by a single link that has cost of the form $k^0 x_t + \sum_r k^r u_t^r$. This can be done through the following algorithm:

1. Consider an OD pair whose routes do not contain other origin or destination nodes in them. Initialize this OD pair as OD_0, where 0 denotes the iteration number. Then, all the routes connecting OD_0 can be replaced by a single link with linear cost function using the algorithm of Appendix A. This link has a cost function that is linear in terms of the flow traveling between OD_0 and the toll values (note that the flow traveling between OD_0 can also constitute demand from other OD pairs apart from the demand of OD_0).

2. In iteration n of the algorithm, we denote the OD pair under consideration as OD_n. Consider an OD pair, denoted as OD_n, that contains OD_{n-1} i.e. if a user is traveling between OD_n then he also travels between the pair OD_{n-1}. Then, OD_{n-1} is connected by a single link that has cost as a linear function of the traffic that flows between the OD_{n-1} (consisting of the demand of OD_{n-1} and demands from other OD pairs). Consider the terms corresponding to the demands of $OD_{n-1}, \ldots, OD_0$ as additional tolls in the cost function of the link joining OD_{n-1} as these do not constitute part of the demand between OD_n. Since Lemmas 8 and 9 ensure that the cost functions of the new links are linear functions of the flows as well as toll values, the single link that replaces all the routes between OD_n has cost that is a linear function of the demands corresponding to $OD_n, OD_{n-1}, \ldots, OD_0$.

3. If the origin and destination nodes of the pair OD_n do not have any incoming and outgoing links, respectively, then stop. Else, go to Step 2.

The proof of correctness of the above algorithm can be proceeded using the following loop invariant: *after iteration n, all the routes joining OD_n can be replaced by a single link that has cost as a linear function of the demands of $OD_n, OD_{n-1}, \ldots, OD_0$.*

To demonstrate how the algorithm works, consider the network in Figure 10 but with two OD pairs: A-C and A-D with demands at time step t being $d_t^1 = \rho x_t$ and $d_t^2 = (1-\rho)x_t$, respectively, where x_t is the total demand at time step t. In the first iteration of the algorithm, A-C is the OD pair under consideration. Then, links 1, 2 and 3 can replaced by a single link 6 that has cost function that linearly varies with the demands of both the OD pairs by the analysis in Appendix A. Denote the cost of link 6 joining nodes A and C at time step t by $k^0(d_t^1 + d_t^2) + k^1 u_t^1 + k^2 u_t^2 + k^3 u_t^3$, where k^0, k^1, k^2 are positive constants and u_t^1, u_t^2, u_t^3 are tolls applied on links $1, 2$ and 3 at time step t, respectively. Alternatively, it can also be considered that link 6 has flow equal to d_t^2 flowing through it at time step t with cost parameter k^0 and toll equal to $k^0 d_t^1 + k^1 u_t^1 + k^2 u_t^2 + k^3 u_t^3$ levied on it. Denote the cost of link 4 at time step t as $cd_t^2 + u_t^4$, where c is a positive constant. Then, links 6 and 4 can be replaced by a single link by Lemma 8 that is linear in terms of the demand values of both the OD pairs and the toll values on different links. Since the demands across individual OD pairs are fixed proportions of the total demand, the single link that connects A and D has cost that is linear in terms of the total demand across the network and the toll values. So, for such networks $TSTT$ is of the same form as before and the solution analysis that we conducted for single OD pair networks carries over.

References

Alden, J. M. and Smith, R. L. (1992). Rolling horizon procedures in nonhomogeneous markov decision processes. *Operations Research*, 40(3-supplement-2):S183–S194.

Altman, E. (1994). Denumerable constrained markov decision processes and finite approximations. *Mathematics of operations research*, 19(1):169–191.

Aviv, Y. and Federgruen, A. (1999). The value iteration method for countable state markov decision processes. *Operations research letters*, 24(5):223–234.

Bertsekas, D. P. (2012a). *Dynamic programming and optimal control.* Athena scientific Belmont, MA.

Bertsekas, D. P. (2012b). Weighted sup-norm contractions in dynamic programming: A review and some new applications. *Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA, Tech. Rep. LIDS-P-2884.*

Bliemer, M., Brederode, L., Wismans, L., and Smits, E.-S. (2012). Quasi-dynamic network loading: adding queuing and spillback to static traffic assignment. In *Transportation Research Board 91st Annual Meeting TRB, Washington DC, United States, 26th January.*

Bowerman, B. L. (1974). Nonstationary markov decision processes and related topics in nonstationary markov chains.

Boyce, D., Janson, B., and Eash, R. (1981). The effect on equilibrium trip assignment of different link congestion functions. *Transportation Research Part A: General*, 15(3):223–232.

Brémaud, P. (2013). *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media.

Cavazos-Cadena, R. (1986). Finite-state approximations for denumerable state discounted markov decision processes. *Applied Mathematics and Optimization*, 14(1):1–26.

Cheng, Q., Liu, Z., Liu, F., and Jia, R. (2017). Urban dynamic congestion pricing: an overview and emerging research needs. *International Journal of Urban Sciences*, 21(sup1):3–18.

Clark, S. and Watling, D. (2005). Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological*, 39(2):119–140.

Correa, J. R., Schulz, A. S., and Stier-Moses, N. E. (2008). A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469.

Cova, T. J. and Johnson, J. P. (2002). Microsimulation of neighborhood evacuations in the urban–wildland interface. *Environment and Planning A*, 34(12):2211–2229.

De Oliveira, L. B. and Camponogara, E. (2010). Multi-agent model predictive control of signaling split in urban traffic networks. *Transportation Research Part C: Emerging Technologies*, 18(1):120–139.

De Palma, A. and Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, 19(6):1377–1399.

Diakaki, C., Papageorgiou, M., and Aboudolas, K. (2002). A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engineering Practice*, 10(2):183–195.

Gao, S. and Chabini, I. (2006). Optimal routing policy problems in stochastic time-dependent networks. *Transportation Research Part B: Methodological*, 40(2):93–122.

Gehlot, H. and Ukkusuri, S. V. (2018). User equilibrium with a policy based link transmission model for stochastic time-dependent traffic networks. *Accepted in Transportmetrica B: Transport Dynamics.*

Ghate, A. and Smith, R. L. (2013). A linear programming approach to nonstationary infinite-horizon markov decision processes. *Operations Research*, 61(2):413–425.

Hasan, S., Schneider, C. M., Ukkusuri, S. V., and González, M. C. (2013a). Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics*, 151(1-2):304–318.

Hasan, S., Zhan, X., and Ukkusuri, S. V. (2013b). Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing*, page 6. ACM.

Hellerstein, D. M. (1991). Using count data models in travel cost analysis with aggregate data. *American journal of agricultural economics*, 73(3):860–866.

Karlin, S. (2014). *A first course in stochastic processes.* Academic press.

Kitamura, R. (1984). Incorporating trip chaining into analysis of destination choice. *Transportation Research Part B: Methodological*, 18(1):67–81.

Liu, D.-K. (1974). *Numerical methods for solving nonlinear equations.* PhD thesis, Texas Tech University.

Meyn, S. P. (1997). The policy iteration algorithm for average reward markov decision processes with general state space. *IEEE Transactions on Automatic Control*, 42(12):1663–1680.

Pi, X. and Qian, Z. S. (2017). A stochastic optimal control approach for real-time traffic routing considering demand uncertainties and travelers choice heterogeneity. *Transportation Research Part B: Methodological*, 104:710–732.

Puterman, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons.

Ramadurai, G., Ukkusuri, S. V., Zhao, J., and Pang, J.-S. (2010). Linear complementarity formulation for single bottleneck model with heterogeneous commuters. *Transportation Research Part B: Methodological*, 44(2):193–214.

Rambha, T. and Boyles, S. D. (2016). Dynamic pricing in discrete time stochastic day-to-day route choice models. *Transportation Research Part B: Methodological*, 92:104–118.

Roughgarden, T. and Tardos, É. (2002). How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259.

Sandholm, W. H. (2002). Evolutionary implementation and congestion pricing. *The Review of Economic Studies*, 69(3):667–689.

Seik, F. T. (2000). An advanced demand management instrument in urban transport: electronic road pricing in singapore. *Cities*, 17(1):33–45.

Sennott, L. I. (1991). Value iteration in countable state average cost markov decision processes with unbounded costs. *Annals of Operations Research*, 28(1):261–271.

Small, K. A. and Yan, J. (2001). The value of "value pricing" of roads: Second-best pricing and product differentiation.

Song, C., Qu, Z., Blumm, N., and Barabási, A.-L. (2010). Limits of predictability in human mobility. *Science*, 327(5968):1018–1021.

Srinivasan, D., Choy, M. C., and Cheu, R. L. (2006). Neural networks for real-time traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, 7(3):261–272.

Steinberg, R. and Zangwill, W. I. (1983). The prevalence of braess' paradox. *Transportation Science*, 17(3):301–318.

Tamor, M. A., Gearhart, C., and Soto, C. (2013). A statistical approach to estimating acceptance of electric vehicles and electrification of personal transportation. *Transportation Research Part C: Emerging Technologies*, 26:125–134.

Tan, Z., Yang, H., and Guo, R.-Y. (2015). Dynamic congestion pricing with day-to-day flow evolution and user heterogeneity. *Transportation Research Part C: Emerging Technologies*, 61:87–105.

Ukkusuri, S. V., Han, L., and Doan, K. (2012). Dynamic user equilibrium with a path based cell transmission model for general traffic networks. *Transportation Research Part B: Methodological*, 46(10):1657–1684.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260.

Viti, F., Catalano, S. F., Li, M., Lindveld, C., and van Zuylen, H. J. (2003). An optimization problem with dynamic route-departure time choice and pricing. In *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*.

Walters, A. A. (1961). The theory and measurement of private and social cost of highway congestion. *Econometrica: Journal of the Econometric Society*, pages 676–699.

Xie, C., Wang, T.-G., Pu, X., and Karoonsoontawong, A. (2017). Path-constrained traffic assignment: Modeling and computing network impacts of stochastic range anxiety. *Transportation Research Part B: Methodological*, 103:136–157.

Yan, X.-Y., Zhao, C., Fan, Y., Di, Z., and Wang, W.-X. (2014). Universal predictability of mobility patterns in cities. *Journal of The Royal Society Interface*, 11(100):20140834.

Zhong, C., Batty, M., Manley, E., Wang, J., Wang, Z., Chen, F., and Schmitt, G. (2016). Variability in regularity: Mining temporal mobility patterns in london, singapore and beijing using smart-card data. *PloS one*, 11(2):e0149222.

Zhu, F. and Ukkusuri, S. V. (2017). Efficient and fair system states in dynamic transportation networks. *Transportation Research Part B: Methodological*, 104:272–289.