Predictive Distance-based Road Pricing — Designing Tolling Zones through Unsupervised Learning

Antonis F. Lentzakis[a,*], Ravi Seshadri[b], Moshe Ben-Akiva[c]

[a] *NCS Hub, NCS Group, 5 Ang Mo Kio Street 62, Singapore 569141*
[b] *DTU Management, Technical University of Denmark, Bygningstorvet, 116, 116A, Kongens Lyngby, Denmark*
[c] *Intelligent Transportation Systems Lab, MIT, 77 Massachusetts Avenue, Room 1-181, Cambridge, M.A., 02139, U.S.A.*

Abstract

Congestion pricing is a standard approach to mitigate traffic congestion in a number of urban networks around the world. The advancement of satellite technology has spurred interest in distance-based congestion pricing schemes, which obviate the need for fixed infrastructure such as gantries that are used in area- and cordon-based pricing. Moreover, distance-based pricing has the potential to more effectively manage traffic congestion. In the context of distance-based congestion pricing, we propose the use of sparse subspace clustering methods employing Elastic Net optimization (SSCEL) and Orthogonal Matching Pursuit (SSCOMP), as well as two hierarchical density-based clustering methods, (OPTICS, HDBSCAN*) for the derivation of tolling zones. These tolling zone derivations are then used within a simulation-based framework for real-time predictive distance-based toll optimization to examine network congestion and performance of the tolling schemes. Within this framework, for a given derivation of tolling zones, tolling function parameters are optimized in real-time using a simulation-based Dynamic Traffic Assignment (DTA) model. Guidance information generation is integrated into the predictive optimization framework and behavioral responses to the information and tolls along dimensions of departure time, route, mode, and trip cancellation are explicitly modeled. For the evaluation of network performance we make use of Travel Speed Index (TSI) data from the real-world Boston Central Business District urban network and demonstrate that tolling zones derived from the sparse subspace cluster-

*Corresponding author: `aflentz@mit.edu`

ing are effective means of operationalizing real-time distance-based toll optimization schemes and can positively impact overall network performance, showing improvements in average travel time and social welfare relative to the baseline.

Keywords: Distance-based Toll Optimization, Sparse Subspace Clustering, Density-based Clustering

---

1. Introduction

Traffic congestion is a serious issue world-wide, which results in large costs to travelers, the environment and economy. Congestion was estimated to result in a total of 5.5 billion hours of time delay and 2.9 billion gallons of fuel expenditure in urban areas in the United States between 2000 and 2010 (Litman, 2019) and the costs of congestion were projected to increase from $121 billion in 2011 to $199 billion in 2020. Mitigating congestion is always a high-priority and also impacts transportation network reliability, driver's comfort and traffic safety. Congestion pricing is a standard approach for congestion mitigation that influences traveler behavior along several dimensions: trip making and frequency, mode, destination, time of day, route, and so on. Traditional approaches to congestion pricing include facility-based and area-based schemes (de Palma and Lindsey, 2011) that rely on physical infrastructure such as gantries or gates for vehicle detection. Unfortunately, the reliance on fixed physical infrastructure makes it difficult to modify or relocate the charging areas or zones. Moreover, these schemes can result in inefficiency in terms of congestion mitigation since they do not differentiate toll charges based on the associated externalities or congestion caused due to differing distances traveled or time spent in congestion. The aforementioned disadvantages of area- and facility-based pricing and the advancement of Global Navigation Satellite Systems (GNSS) have focused attention on usage-based tolling wherein toll charges depend on the distance-traveled or the time spent in congestion (see Smith et al. (1994) and Bonsall and Palmer (1997) for a detailed discussion on the comparative performance of distance- and time-based schemes). Singapore is in the process of transitioning to such a GNSS-based electronic road-pricing scheme (ERP2) (LTA, 2016, 2021). Distance-based schemes may be operationalized by dividing the urban area into zones and charging a distance-based toll such that the tariff varies across zones and by time-of-day. The motivation for the use of tolling zones (instead of a single distance-based scheme over the entire network) is that it provides the flexibility to adjust the tolling rates based on road-type and congestion levels, thereby improving overall efficiency gains.

Past research on area and cordon-based real-time toll optimization has typically applied reactive approaches (where the optimization of tolls is not based on forecasts of future traffic conditions, but rather on prevailing traffic conditions) for small corridor networks and there are few studies that adopt a predictive approach in the context of large networks (Gupta et al., 2016, 2020). A more detailed discussion of cordon and area-based real-time toll optimization may be found in Gupta et al. (2020). As noted previously, in contrast with cordon- and area-based schemes, distance-based tolling schemes involve partitioning the network into zones, and levying a toll within each zone that is a function of distance traveled (linear toll functions are considered in Gu et al. (2018); Yang et al. (2012); Zhu and Ukkusuri (2015), and piece-wise linear functions are used in Liu et al. (2014); Meng et al. (2012); Sun et al. (2016)). Distance-based toll optimization problems have largely been formulated as simulation-based optimization problems (Gu and Saberi, 2019b; Gu et al., 2018; Lentzakis et al., 2020), non-linear programs (Yang et al., 2012) and mathematical programs with equilibrium constraints or MPEC (Liu et al., 2014; Meng et al., 2012), which are solved by global optimization approaches (Liu et al., 2014), meta heuristics (Lentzakis et al., 2020; Meng et al., 2012), reinforcement learning (Zhu and Ukkusuri, 2015) and feedback controllers (Gu and Saberi, 2019b; Gu et al., 2018). With the exception of Lentzakis et al. (2020), these approaches are based on prevailing network conditions (i.e., they are reactive as opposed to proactive), and do not consider elastic demand or the integration of guidance information generation.

Several studies have also examined the partitioning of networks utilizing flow, speed and density data (Ji and Geroliminis, 2012; Lentzakis et al., 2014; Saeedmanesh and Geroliminis, 2017) for the design of traffic management schemes utilizing the Network Fundamental Diagram (NFD) concept. Although area- and cordon-based pricing has been studied in great detail (Geroliminis and Levinson, 2009; Simoni et al., 2015; Zheng et al., 2016, 2012), distance-based pricing in particular has only recently received attention on idealized networks (Daganzo and Lehe, 2015), using nested regions (Gu et al., 2018) and at the link-level (Simoni et al., 2019). With the exception of Lentzakis et al. (2020), there has been limited research on systematic approaches for the derivation of tolling zones within distance-based toll optimization strategies. Due to the increasing significance of distance-based road pricing in traffic network management and operations, this paper addresses the problem of how to define tolling zones and proposes the application of sparse subspace clustering methods to define parsimonious sets of tolling zones. The performance of these methods is evaluated within a framework for real-time toll optimization which generates predictive optimized distance-based toll strategies combined with guidance information. This paper contributes to the existing literature in the following respects:

1. We apply sparse subspace and hierarchical density-based clustering methods for the derivation of tolling zones that utilize location coordinates and travel speed indices (TSI) as features. The key advantage of using sparse subspace clustering techniques is that they enable the effective use of high-dimensional temporal network performance data (for example, travel speeds at a resolution of five minutes) directly in the clustering algorithm. This provides a potentially promising alternative to the procedure proposed in Lentzakis et al. (2020) where the clustering algorithm is applied to a single aggregate measure of network performance (over the entire peak period) for each link. In this paper, we focus specifically at the performance of the different clustering algorithms and the implications for toll design/policy.
2. The proposed clustering methods are evaluated using a framework for real-time distance-based predictive toll optimization on the Boston CBD network and yield insights into their performance and suitability for deployment wherein one of our primary goals is minimization of computational effort.

2. Framework for Predictive Distance-based Toll Optimization

In this section, we summarize the real-time distance-based predictive toll optimization framework (more details may be found in Lentzakis et al. (2020)), the optimization problem formulation, the proposed clustering methods for tolling zone derivation and the algorithmic solution for the optimization problem.

2.1. Framework

The framework, shown in Figure 1, uses DynaMIT2.0 - a simulation-based Dynamic Traffic Assignment (DTA) system developed at the MIT Intelligent Transportation Systems Lab (Ben-Akiva et al.,

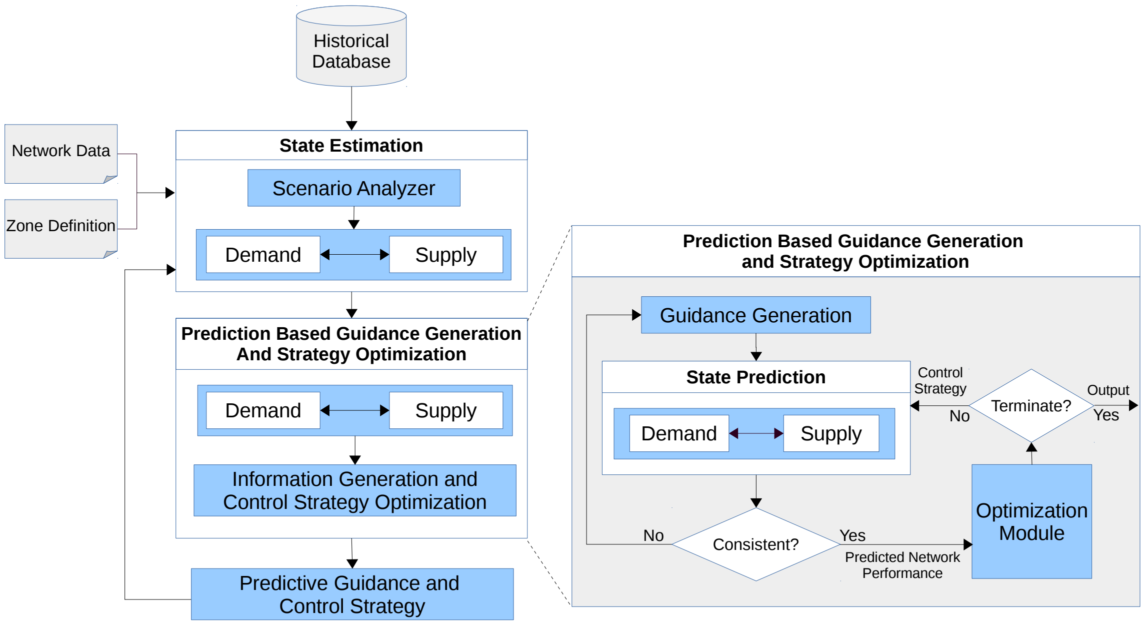


Figure 1: Real-time distance-based predictive toll optimization framework

2010; Lu et al., 2015b). DynaMIT2.0 employs a rolling horizon approach involving two key modules, state estimation and state prediction. The state estimation process uses a combination of historical data, real-time traffic surveillance data, and prevailing network control strategies (such as distance-based toll optimization) to estimate the current state of the network. It used detailed models of demand (pre-trip models of departure time, route and mode choice), supply (mesoscopic traffic simulator that combines speed-density relationships and a deterministic queuing model) and their interactions. Following this, the state prediction module generates forecasts of traffic conditions for a pre-specified prediction horizon (origin-destination demands and supply parameters are forecasted for the future using an autoregressive process). The strategy optimization and guidance generation modules in conjunction use the state predictions to first, optimize control strategies for the prediction horizon and second, generate guidance information (traveler information) for the prediction horizon. The evaluation of candidate control strategies makes use of network predictions and guidance information that are consistent, i.e., the guidance information is as close as possible to actual predicted network travel times (see Figure 1 and Ben-Akiva et al. (2010) for more on this aspect of consistency).

3. Problem Formulation and Solution

In this section, we describe the optimization problem formulation (based on the framework described in Section 2) including details of the demand model within the DTA system, and the solution approach.

3.1. Context and Tolling Function Definition

We represent the transportation network of interest as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ denotes the set of n network nodes and $\mathcal{A}$ denotes the set of m links. The network is partitioned into $l = 1 \dots L$ tolling zones, where every zone l is defined by a subset of network links $\mathcal{A}_l \subseteq \mathcal{A}$. For each zone l, we define a tolling function $\phi_l(\boldsymbol{\theta_l^t}, D_l)$ that maps distance traveled within the zone l, D_l to the toll amount; $\boldsymbol{\theta_l^t}$ is a vector of parameters that defines the tolling function in time interval t. Further, it is assumed that the toll payable in a zone is bounded, i.e $\tau_{LB} \leq \phi_l(\boldsymbol{\theta_l^t}, D_l) \leq \tau_{UB}, \forall l = 1, 2..., L\ \forall t = 1, 2, ..., T$.

Denote the length of the state estimation interval in DynaMIT2.0 by Δ (usually 5 minutes) and assume that the prediction horizon is composed of H such intervals so that the size of the prediction horizon is $H\Delta$. We assume that the prediction horizon and the optimization horizon are identical. Further, the tolling function parameters do not vary within a given time interval of size Δ and these tolling intervals coincide with DynaMIT2.0 estimation intervals. For an arbitrary estimation interval $[t_0 - \Delta, t_0]$, let $\boldsymbol{\theta^h} = (\boldsymbol{\theta_1^h}, \boldsymbol{\theta_2^h} ... \boldsymbol{\theta_L^h})$ represent the vector of tolling function parameters for the time period $[t_0 + (h-1)\Delta, t_0 + h\Delta]$ where $h = 1, ..., H$. Accordingly, for the current optimization horizon, the decision variables are $\boldsymbol{\theta} = (\boldsymbol{\theta^1}, \boldsymbol{\theta^2}, ..., \boldsymbol{\theta^H})$.

Implementing a system with complex zone-based tariffs that vary every five minutes is likely to impose unreasonable burdens on drivers that may compromise acceptability of the system. An added issue is that drivers may not have a viable alternative if for example they suddenly find themselves entering a zone where the tariff has increased substantially. Hence, we assume that drivers are charged the predicted toll that the system provides to them at the point of departure (more precisely, the point at which they make their decision, which may be up to 15 minutes prior to their actual departure). The underlying premise – justifiable given our rolling horizon design – is that the predictions of the toll in the future do not deviate appreciably from the actual implemented tolls. The rolling horizon framework is demonstrated in Figure 2

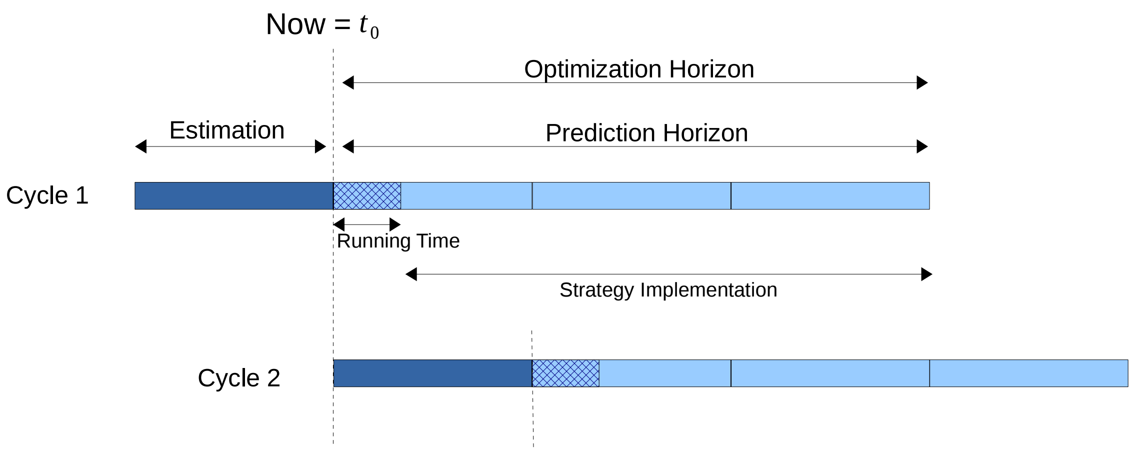


Figure 2: Rolling Horizon Approach for Tolling Function Optimization

Consider the set of vehicles $v = 1, ..., V$ that are on the network during the prediction horizon $[t_0, t_0 + H\Delta]$. For each vehicle v, we denote the experienced trip travel on its chosen route by tt^v and the predictive guidance information by $\mathbf{tt}^g = (\mathbf{tt_i^g}; \forall i \in A)$, where $\mathbf{tt_i^g}$ represents a vector of time dependent travel times for link i . Note that the vehicle travel times $\mathbf{tt} = (tt^v; v = 1, ..., V)$ are obtained from the state prediction module of DynaMIT2.0, which we characterize through a single constraint that represents the coupled demand and supply simulators as:

$$G(\mathbf{x^p}, \gamma^\mathbf{p}, \mathbf{tt^g}, \boldsymbol{\theta}) = \mathbf{tt} \tag{1}$$

Where $\mathbf{x^p}$, $\gamma^\mathbf{p}$ represent the forecasted demand and supply parameters for the prediction horizon, and

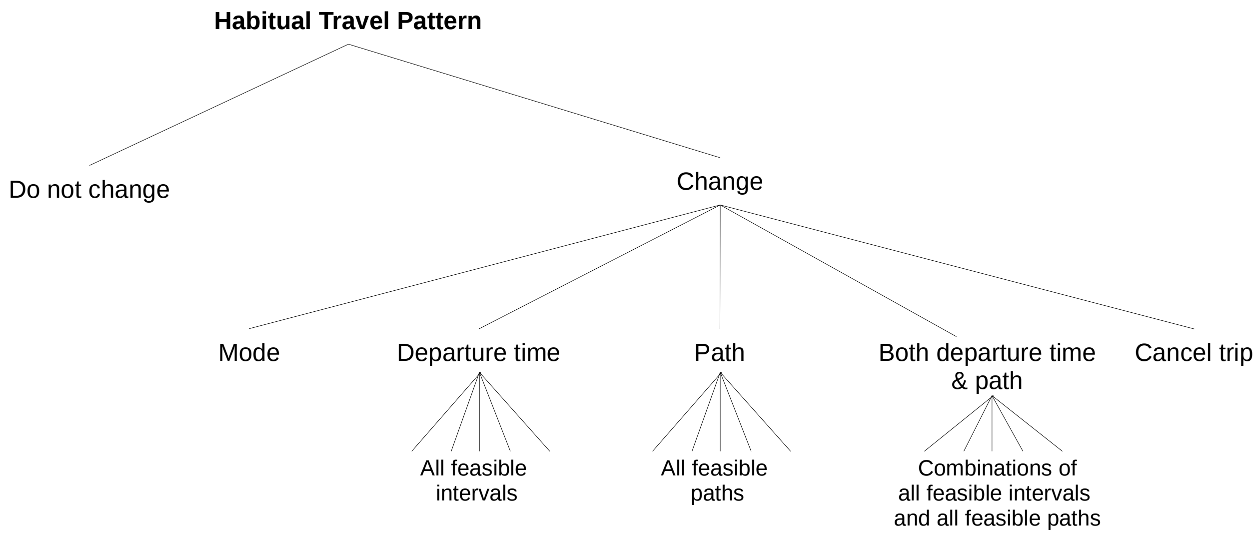


Figure 3: Pre-trip Behavior Model

$\boldsymbol{\theta}$ is the vector of tolling function parameters. As noted previously, the state prediction module ensures consistency between $\mathbf{tt^g}$ and $\mathbf{tt}$.

3.2. Pre-trip Behavioral Model with Elastic Demand

The pre-trip response of users to the travel time guidance and distance-based tolls is modeled using a path-size nested logit model with heterogeneous value of time (illustrated in Figure 3) that captures decisions of mode choice, trip cancellation, departure time and path (notation is provided in Table 1). We provide a brief description of the model here, for completeness (more details may be found in Lentzakis et al. (2020)).

In response to pre-trip information and tolls, a traveler may alter his/her habitual travel pattern, which may include changing mode, canceling trip, changing departure time or path, or changing departure time and path. This results in elastic total demand w.r.t. traffic congestion. The options of mode modeled are private car (drive alone) and public transit. The utility of change to transit for vehicle v is given by:

Table 1: Pre-trip Model - Abbreviations

Abbreviation	**Variable**
β_{CM}	Alternative Specific Constant (ASC) for change of mode to transit
β_{CT}	ASC for canceling trip
β_{CDT_d}	ASC for departure in time interval d
c_m^v	monetary cost for traveling with non-private (transit) mode
c_{dp}^v	toll charge for departure via path p in interval d
c_p^v	toll for switching to path p
t_m^v	travel time associated with non-private (transit) mode
t_{dp}^g	travel time (guidance) for departure via path p in interval d
t_p^g	travel time (guidance) for switching to path p
$at_{d'p'}^{hab}$	arrival time (habitual)
at_{dp}^g	arrival time (predicted) for departure via path p in time interval d
β_c^v	monetary cost coefficient
β_t^v	travel time coefficient
β_E	schedule delay early coefficient
β_L	schedule delay late coefficient
PS_p	path size variable
C_*	utility relating to number of left turns/signalized intersections and path length
ε_*	error component

$$U^v(CM) = \beta_{CM} + \beta_c^v c_m^v + \beta_t^v t_m^v + \varepsilon_m \tag{2}$$

The utility of departing at time interval d and choosing path p for vehicle v is given by:

$$\begin{aligned} U_{dp}^v = \beta_{CDT_{dp}} &+ \beta_c^v c_{dp}^v + \beta_t^v t_{dp}^g \\ &+ \beta_E \max(at_{d'p'}^{hab} - at_{dp}^g, 0) + \beta_L \max(at_{dp}^g - at_{d'p'}^{hab}, 0) \\ &+ log(PS_p) + C_{dp} + \varepsilon_{dp} \\ where: & \\ c_{dp}^v = \sum_{l=1}^{L} &\phi_l(\boldsymbol{\theta_l^{t_{v,l}}}, D_l^v), \end{aligned} \tag{3}$$

and $t_{v,l}$, D_l^v denote the predicted time of entry of vehicle v into zone l and the total distance traveled by vehicle v in zone l, respectively. Note that if there are a total of N combinations of path and departure time choices in the choice set, the alternative specific constant β_{CDT_d} can only appear in $(N-1)$ utilities. The utility of canceling trip altogether is given by:

$$U^v(CT) = \beta_{CT} + \varepsilon_{CT} \tag{4}$$

Thus, the probability of vehicle v choosing alternative c within the choice set C is given by:

$$P^v(c|C) = \frac{e^{\mu V_c^v}}{\sum_{a \in C} e^{\mu V_a^v}} \tag{5}$$

where V_c^v is the systematic utility given by $V_c^v = U_c^v - \varepsilon_c$ and μ is a scale parameter . The en-route choice model defines response of users in terms of path-choice to the toll and predictive travel time guidance. It is also formulated as a multinomial path size logit model where the utility of switching to path p is given by:

$$
\begin{aligned}
\cup_p^v &= \beta_c^v(tc_p^v) + \beta_t^v(tt_p^g) + log(PS_p) + C_p + \varepsilon_p \\
where &: \\
tc_p^v &= \sum_{l=1}^{L} \phi_l(\boldsymbol{\theta_l^{t_{v,l}}}, D_l^v)
\end{aligned}
\tag{6}
$$

Note that owing to the design of the distance-based tolls, which require that users are charged upfront at the beginning of a trip, we assume that en-route changes to the path are not made.

3.3. Optimization Formulation

The objective function for the toll optimization problem, formulated from the standpoint of the traffic regulator, is total social welfare (SW), which is the sum of the consumer surplus and the producer surplus. In this context, the consumer surplus (CS) is defined as the sum of the experienced utilities across all travelers, derived at the end of each simulation run, and the producer surplus is the net revenue, denoted by TP, which is simply the toll revenue minus variable costs (fixed costs are ignored), $TP = TR - VC$. We assume that the variable costs are a proportion of the toll revenue (the proportionality factor is denoted by $\alpha < 1$). Thus, the social welfare is given by:

$$
\begin{aligned}
SW &= CS + TP \\
&= CS + (TR - VC) \\
&= \sum_{v=1}^{V} \frac{U^v}{|\beta_c^v|} + \left[(1-\alpha) \times \sum_{v=1}^{V} c^v\right]
\end{aligned}
\tag{7}
$$

The absolute value of β_c^v is used to translate CS into dollar equivalents. The distance-based toll optimization problem is formulated as a simulation-based optimization problem in Equation 8, where the objective is social welfare, the decision variables are the vector of tolling function parameters for the current optimization horizon, and the constraints are toll bounds and the DTA model system.

$$
\begin{aligned}
&\max_{\boldsymbol{\theta}} \left[\sum_{v=1}^{V} \frac{U^v}{|\beta_c^v|} + (1-\alpha) \times \sum_{v=1}^{V} c^v \right] \\
&\textbf{s.t.} \\
&G(\mathbf{x^p}, \gamma^{\mathbf{p}}, \mathbf{tt^g}, \boldsymbol{\theta}) = \mathbf{tt} \\
&\tau_{LB} \leq \phi_l(\boldsymbol{\theta_l^h}, D_l^v) \leq \tau_{UB}, \forall v = 1, 2, ..., V; l = 1, 2..., L; h = 1, 2, ..., H
\end{aligned}
\tag{8}
$$

The upper and lower bounds τ_{LB}, τ_{LB} are imposed to allow for tolling function values within a safe and acceptable range, suitable for real-life implementations.

3.4. Solution Algorithm

Due to the highly non-convex nature of the objective function in 8, we apply a real-coded Genetic Algorithm (GA) to solve the optimization problem in 8. More details on the GA algorithm may be found in Lentzakis et al. (2020). Computational performance is enhanced by utilizing parallelization wherein the evaluations of different candidate solutions within an iteration of the GA are performed in parallel.

4. Tolling Zone Design through Unsupervised Learning

For most tolling-related implementation decisions, at least in the USA, tolling zone boundaries are subject to extreme political scrutiny, Environmental Justice reviews, exemptions for residents of certain areas, etc. Unsupervised machine learning utilizes historical data to reveal patterns, similarities or hidden structure and can contribute to changing the current state-of-affairs, with regards to tolling zone design. In this paper, we posit that it would be beneficial for a city or highway operator to rely on unsupervised learning approach in any sort of real-world setting, for expeditious implementation. Even in the case that stakeholder involvement is mandatory, tolling zone derivation through unsupervised learning can significantly augment the decision-making process. Unsupervised learning can be approached through different techniques such as clustering, association rules, and dimensionality reduction. Our focus will be on clustering. One of the inputs with significant impact on our distance-based tolling system performance is the tolling zone derivation. This input specifies which links belong to each tolling zone and the number of zones. Each tolling function $\phi_l(\boldsymbol{\theta_l^h}, D_l^v)$ corresponds to one tolling zone. Past literature for partitioning urban traffic networks used datasets based on speed, flow, density (Gu and Saberi (2019a); Ji and Geroliminis (2012); Lentzakis et al. (2014); Saeedmanesh and Geroliminis (2017)) and, more recently, marginal cost toll data (Lentzakis et al., 2020). In our case, the travel speed index (TSI) is used, a widely used quantitative indicator that employs link speed normalization (Li and Xiao, 2020), given the fact that identical link speed levels might reflect different traffic conditions. Speed information for toll setting is currently used in a similar fashion as in Singapore's ERP system, (Lehe, 2019). It should be noted that the decision to use travel speed indices, rather than marginal cost tolls (MCT) used in Lentzakis et al. (2020), has to do with the fact that, in this work, one of our main goals was to reduce computational effort, both during data preprocessing and the predictive distance-based toll optimization framework implementation, placing real-world applicability at the forefront. Should circumstances allow it, the possibility of using MCT as a feature should definitely be explored.

4.1. Clustering Approaches

Elhamifar and Vidal (2009), inspired by compressed sensing (Lee et al., 2007), introduced Sparse Subspace Clustering (SSC), which makes use of the self-expressiveness property to construct the affinity matrix (which quantifies the extent of pairwise similarity between a set of data points). Self-expressiveness (Elhamifar and Vidal, 2013) describes the fact that a data point found in a union of subspaces can be represented as the linear combination of other data points. Based on the computed affinity matrix, spectral clustering is applied to derive the underlying subspaces. While subspace clustering methods have been used extensively for, among others, temporal video segmentation and switched system identification (Bako, 2011; Rao et al., 2009), only recently, has this technique come to the attention of the transportation research community. Zhang et al. (2019) employed SSC to classify spatiotemporal taxi patterns with regards to their passenger searching behavior. For our experiments we selected to compare two Sparse Subspace Clustering variants, SSCEL and SSCOMP, employing

Elastic Net optimization (You et al., 2016a) and Orthogonal Matching Pursuit (You et al., 2016b) respectively, against two well-known hierarchical density-based clustering methods, OPTICS (Ankerst et al., 1999), (Ordering Points To Identify the Clustering Structure), and HDBSCAN* (Campello et al., 2013), (Hierarchical Density-Based Spatial Clustering of Applications with Noise).

4.1.1. Sparse Subspace Clustering Methods

For the Sparse Subspace Clustering application, we selected two variants,SSCEL and SSCOMP. We exploited the property of self-representation to learn the affinity matrix, to be subsequently used in our implementation of spectral clustering. As noted previously, data self-expressiveness (Elhamifar and Vidal, 2013) describes the fact that a data point found in a union of subspaces can be represented as the linear combination of other data points, expressed through the following optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{C}} \|\mathbf{C}\|_1 \\ &\textbf{s.t.} \\ &\mathbf{X} = \mathbf{XC} \\ &diag(\mathbf{C}) = 0 \end{aligned} \tag{9}$$

Where $\mathbf{X} \in \mathbf{R}^{D \times N}$ is the data point matrix and $\mathbf{C} \in \mathbf{R}^{N \times N}$ is the self-expression coefficient matrix. In practice, however, solving N such problems over N variables may be computationally expensive for large N. Instead, the optimization problem is expressed as follows:

$$\begin{aligned} &\min_{\boldsymbol{c}_j} \|\mathbf{x}_j - \mathbf{X}\mathbf{c}_j\|_2^2 \\ &\textbf{s.t.} \\ &\|\mathbf{c}_j\|_0 \leq \mathbf{k} \\ &diag(\mathbf{C}) = 0 \end{aligned} \tag{10}$$

We can now efficiently solve the above problem using the Orthogonal Matching Pursuit algorithm, as described in You et al. (2016b). Orthogonal Matching Pursuit selects a single column of $\mathbf{X}$ each time,

$\mathbf{x}_j$, such that the absolute value of the dot product with the residual $\mathbf{c}_j$ is maximized and the coefficients are computed until k columns are selected. Subsequently, we learn the affinity matrix $\mathbf{W}$ through data self-representation as: $\mathbf{W} = |\mathbf{C}| + |\mathbf{C}^T|$. Alternatively, we may employ Elastic Net regularization for scalable subspace clustering. Following You et al. (2016a), we used an active set algorithm that efficiently solves the elastic net regularization subproblem, which follows below, by capitalizing on the geometric structure of the elastic net solution:

$$\min_{\boldsymbol{c}_j} \lambda \|\mathbf{c}_j\|_1 + \frac{1-\lambda}{2}\|\mathbf{c}_j\|_2^2 + \frac{\gamma}{2}\|\mathbf{x}_j - \mathbf{X}\mathbf{c}_j\|_2^2 \tag{11}$$

Where $\lambda \in (0, 1]$ and $\gamma > 0$. In the majority of solution approaches for the Subspace Clustering problem, after learning the affinity matrix, spectral clustering is applied to the resulting matrix to derive the final clustering.

4.1.2. Hierarchical Density-based Clustering Methods

Hierarchical density-based clustering methods are gaining traction among the research community, exhibiting robustness during parameter selection and being able to cope with clusters characterized by large inter-cluster density variability, unlike their non-hierarchical predecessor, DBSCAN (Schubert et al., 2017).

OPTICS utilizes hyperparameters ϵ and κ, representing the maximum ball radius with each data point at its center and the minimum density threshold, respectively. Assuming a metric space (X, d) comprising of a set of data points $X = \{x_1, x_2, ..., x_n\}$, a data point x is considered to be a core point with respect to ϵ and κ if its ϵ-neighborhood $N_\epsilon(x)$ contains a minimum of κ data points. Two core points x_i, x_j are ϵ-reachable with respect to ϵ and κ if they are both contained within each others ϵ-neighborhood. Two core points x_i, x_j are density-connected with respect to ϵ and κ if they are directly or transitively ϵ-reachable. A cluster is the largest possible group of data points, where each two points are considered connected in terms of density. In OPTICS data points are assigned a core distance $d_{\text{core}}^{\epsilon,\kappa}(x)$ to the κ-th nearest neighbor, for varying degrees of density. The reachability-distance $d_{\text{reach}}^{\epsilon,\kappa}(x_i, x_j)$ is the maximum between the core distance of x_i and the distance between data points x_i,x_j. A single global ϵ' value is used to extract a flat clustering.

HDBSCAN* is similar to OPTICS with parameter $\epsilon = \infty$ and a different technique, based on cluster stability, is utilized for flat clustering. In the case of HDBSCAN*, we have $d^{\kappa}_{\text{core}}(x_i)$ representing the κ-th nearest neighbor distance. For a fixed κ and a range of possible ϵ values, the mutual reachability distance $d^{\kappa}_{\text{mreach}}(x_i, x_j)$ is used to generate a complete hierarchy of clusterings. Thus, for any fixed ϵ value, the clustering produced by DBSCAN at a given level in the hierarchy is the clustering obtained for the corresponding ϵ value.

The selected hierarchical density-based clustering methods result in clusterings where some data points are considered noise. A feasible derivation of tolling zones, however, must involve the assignment of all data points to clusters. In order to address this issue, we perform a secondary assignment where all noise data points are assigned to the closest clusters (using Euclidean distance).

4.2. Clustering Performance Metrics

As clustering performance metrics, the Silhouette Coefficient (SC) (Rousseeuw, 1987) and the Davies-Bouldin index (DB) (Davies and Bouldin, 1979) were selected. SC is the average for the entire dataset of the silhouette, which measures cohesion and separation for each cluster and ranges from $[-1, 1]$, where -1 represents an inappropriate clustering (within-cluster variability is large and between cluster variability is small), 0 represents overlapping clusters and 1 represents highly dense clustering. DB is a function of the ratio of intra-cluster scatter to inter-cluster separation. DB values closer to 0 indicate a better clustering result.

4.3. Clustering Results

It would be preferable that the selected clustering methods produce clustering results that are of high quality, according to our previously presented internal evaluation indices, but also do not preclude any sort of practical application, due to high computational cost, incurred on the distance-based toll optimization framework. This translates into a static tolling zone derivation (non-varying during the simulation), with a reasonably low number of tolling zones. For our dataset, besides spatial coordinates, we decided to use the travel speed index (TSI) for each link as an additional feature, calculated as follows:

$$\mathrm{TSI}_i = 1 - \frac{\nu_i}{\nu_i^f} \tag{12}$$

Where ν_i, ν_i^f the link speed and free flow speed for link i respectively. Simulated speed data at the segment and link level, obtained from a calibrated DynaMIT2.0 model of the Boston CBD (Lu et al., 2015a), were used to derive the tolling zone derivations. In the case of static partitioning schemes derived offline, a preferable alternative to using the average of TSI across specific intervals, as is the case for our hierarchical density-based clustering approaches, would be to use the TSI values for all time intervals, i.e., the entirety of our dataset, since self-representation, an integral part of sparse subspace clustering, is amenable for use of datasets with spatiotemporal attributes (Hashemi and Vikalo, 2018; Pham et al., 2012). The Boston CBD network, shown in Figure 4, has 846 nodes, 1746 links, 3085 segments, 5057 lanes and 13080 Origin-Destination pairs.

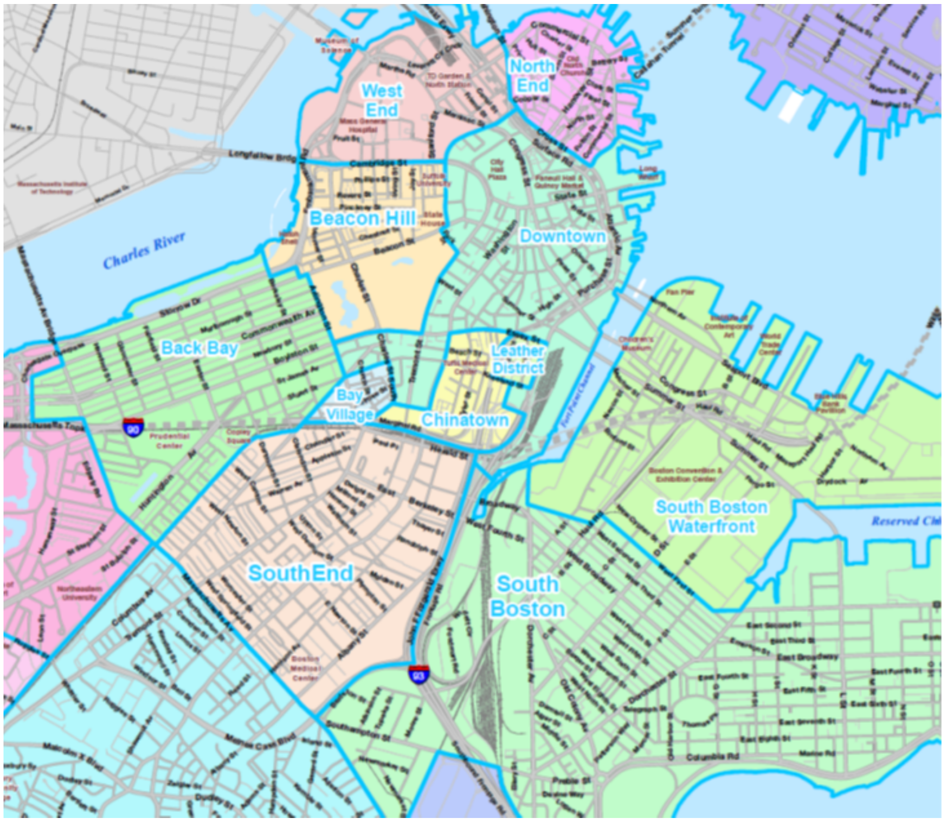
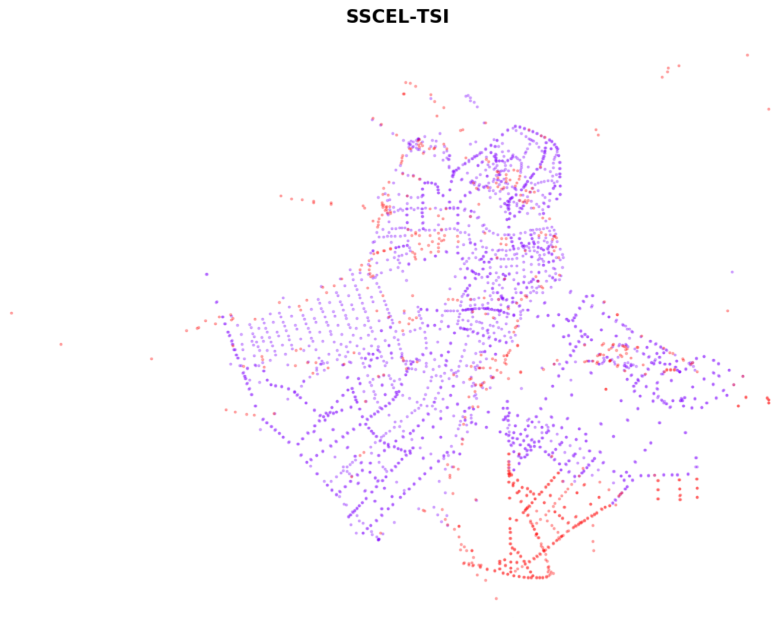
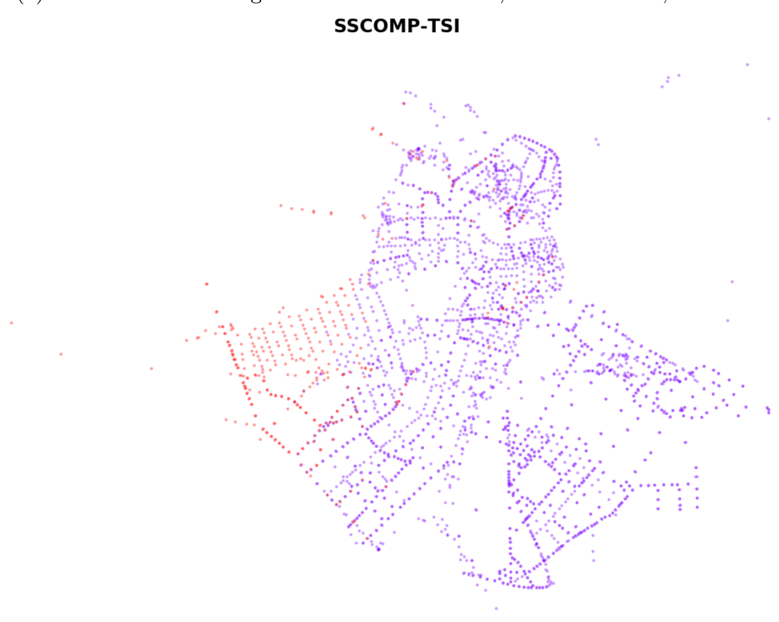


Figure 4: Boston CBD Network

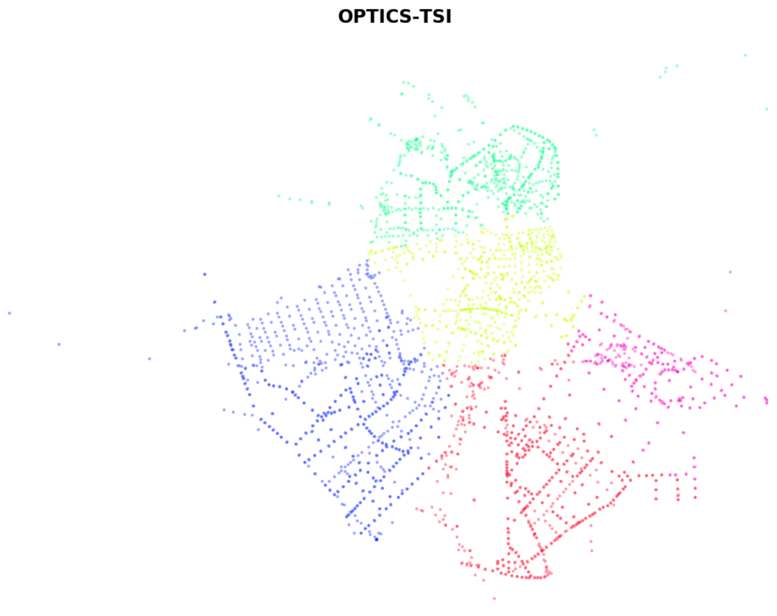


(a) 2 zones derived using SCCEL for feature TSI, with SC=0.114, DB=3.925

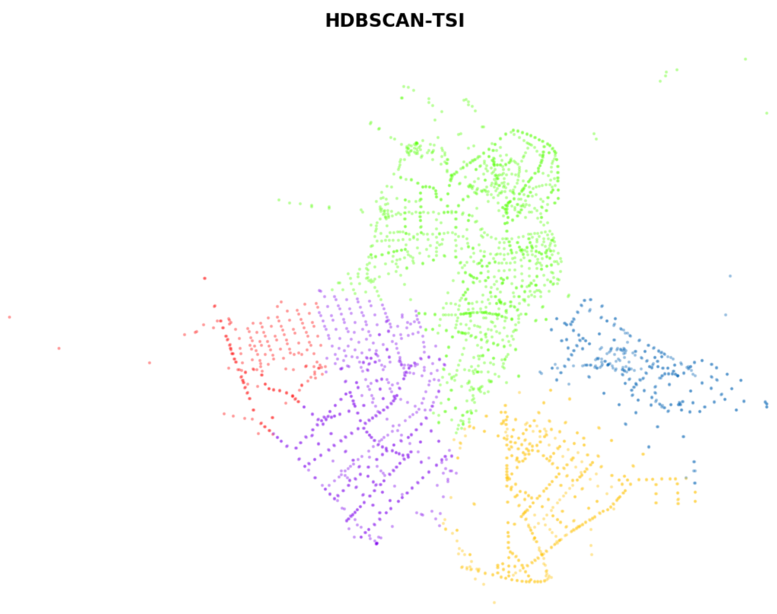


(b) 2 zones derived using SSCOMP for feature TSI, with SC=0.198, DB=1.276

Figure 5: Clustering results and tolling zones (Sparse Subspace Clustering)



(a) 5 zones derived using OPTICS for feature TSI, with SC=0.387, DB=0.858



(b) 5 zones derived using HDBSCAN* for feature TSI, with SC=0.400, DB=0.776

Figure 6: Clustering results and Tolling zones (Hierarchical Density-based Clustering)

5. Experiments: Boston CBD Network

5.1. Experimental Design

In order to investigate the impact tolling zone derivations –which are derived from unsupervised learning methods– have on the performance of adaptive distance-based congestion pricing schemes, when applied on an urban network, experiments are conducted on the Boston CBD network illustrated in Figure 4. A linear tolling function is considered with lower and upper bounds on the toll charged in each zone (i.e $\phi_l(\boldsymbol{\theta_l^t}, D_l) = \theta_{l1}^t + \theta_{l2}^t D_l$; and $0 \leq \phi_l(\boldsymbol{\theta_l^t}, D_l) \leq 1.5$). The simulation period is from 06:00-09:00 covering the morning peak. As noted earlier, historical demand and supply parameters are obtained from prior offline calibrations of DynaMIT2.0 for the Boston Central Business District network (Azevedo et al., 2018). The estimation interval is 5 minutes and the prediction horizon is 30 minutes. The Boston network we consider contains 846 nodes, 1746 links, 3085 segments, 5057 lanes and 13080 origin-destination pairs.

The performance measures are calculated for the population of vehicles with habitual departure time within 06:00-09:00 (these drivers may later change the departure time in response to the traffic conditions). A *warm-up* period of 15 min is used and the last 15 min of the simulation is a *cool-down* period without toll optimization to ensure that all the vehicles with habitual departure time in 06:00-09:00 finish their trips.

The mean and standard deviation of the value of time are S$23.5 and S$5.75 respectively. The cost coefficient for each vehicle is calculated from the lognormally distributed sampled value of time. The parameters of the pre-trip choice model are summarized in Table 2.

A total of six scenarios are considered, which are summarized in Table 3. All scenarios involve dynamic tolls computed using the framework described in Section 2, with the exception of the base scenario (**B0**) which is the *No Toll* case. Recall that the state estimation interval is 5 minutes implying that in the case of distance-based pricing schemes the tolling function parameters vary every 5 minutes. The simulations were run using Ubuntu Linux on an HPC Cluster, with 5x60 cores and 5x250GB RAM.

The base scenario **B0** was calibrated to replicate prevailing traffic conditions in the Boston CBD

Table 2: Pre-trip Behavioral Model - Parameters

Parameter	**Value**
β_{CM}	-0.5
β_{CT}	-12
β_{CDT_1}	-0.12
β_{CDT_2}	-0.79
β_{CDT_3}	-1.15
β_{CDT_4}	- 1.65
β_t^v	-0.008
β_E	-0.004
β_L	-0.016

(refer to Azevedo et al. (2018) for more details). However, given the fact that a No Toll base scenario may not provide specific information regarding what portion of the performance uplift stems from these novel distance-based tolling schemes, rather than the inherent effects of using tolling to internalize the congestion externality, we are also considering a comparison scenario **B1**, (termed **UNIREG-TSI**), which employs predictive distance-based tolling on the the Boston CBD network as a unitary region. Scenarios **B2, B3, B4, B5** employ predictive distance-based tolling and differ only in the derivation of the tolling zones. In scenario **B2** (termed **SSCEL-TSI**), tolling zones are defined based on TSI data using SSCEL. In scenario **B3** (termed **SSCOMP-TSI**), tolling zones are defined based on TSI data using SSCOMP. In scenario **B4** (termed **OPTICS-TSI**), tolling zones are defined based on TSI data using OPTICS, and finally, in scenario **B5** (termed **HDBSCAN*-TSI**) they are based on TSI data using HDBSCAN*. Scenarios **B0-B5** are evaluated on three performance measures, total social welfare (SW), consumer surplus (CS) and average travel time (TT) to capture overall societal benefits, together with the impact on individual travelers.

Table 3: Simulation scenarios

Scenario	Tolling Scheme	Description
B0	No Toll	*No tolling scheme in place*
B1	Predictive distance-based	*Tolling zone encompassing entire network: UNIREG-TSI*
B2	Predictive distance-based	*Tolling zones derived from: SSCEL-TSI*
B3	Predictive distance-based	*Tolling zones derived from: SSCOMP-TSI*
B4	Predictive distance-based	*Tolling zones derived from: OPTICS-TSI*
B5	Predictive distance-based	*Tolling zones derived from: HDBSCAN*-TSI*

5.2. Results

The performance measures for all simulation scenarios are summarized in Table 4, the differences in SW and CS (in $ amounts) of scenarios **B1-B5** relative to the base scenario B0 are presented in Figure 7a, and the relative performance in terms of average travel time (% improvement) over the base scenario B0 is illustrated in Figure 7b. From Table 4, **B1-B5** exhibit an increase between $182623.5 - $206866.5 and $64814.9 - $127062.4, for SW and CS respectively, relative to **B0**. The average SW gain per traveller, relative to the No Toll case is found to be around $1.69 for those acquired via sparse subspace clustering and around $2.15 for tolling zone derivations acquired via hierarchical density-based clustering.

Observe that all the scenarios yield a positive consumer surplus indicating that net user benefits are positive even prior to any use of the toll revenues. This is a surprising finding and in contrast with several past studies that have estimated negative user benefits (for example Eliasson and Mattsson (2006) and De Palma et al. (2005)). We conjecture that this is a result of several factors. First, as noted by Van Den Berg and Verhoef (2011)), in the case when there is heterogeneity in the value of time (and values of schedule delay), the net user benefits may depend in large part on the extent and nature of heterogeneity. In experiments on a variant of the standard bottleneck model including departure time choice and a transit alternative (with heterogeneity in value of time, schedule delay, early and late), Chen (2022) find that when the coefficient of variation in the value of time exceeds around 0.5, the net user benefits start to become positive (even before accounting for distribution of toll revenues). Second, our system integrates the provision of consistent guidance information with the optimization of tolls. These two factors coupled with the high levels of initial congestion may be the reason why we observe positive net user benefits even prior to a redistribution of toll revenues. Similar tests across different network topologies and spatio-temporal congestion patterns are required to determine whether this finding is a peculiarity of our context and network. The significant variation of differences in both welfare and CS across the five schemes confirms that the performance of distance-based tolling schemes is appreciably affected by the definition of the tolling zones. First, observe that the two sparse subspace clustering approaches yield quite different clusters and varying outcomes in terms of both CS and welfare. Scenario **B2** (SCCEL) yields the second lowest overall welfare, which is only marginally higher than scenario **B1** where the entire network is treated as a single zone. The reason for the relatively poor performance is two-fold. First, as is apparent in Figure 5a, SCCEL results in clusters that lack spatial compactness. In other words, zones are 'non-contiguous' and links in different parts of the network belong to the same zone (links belonging to the red cluster or zone in particular). This clearly poses an issue in the toll optimization, since the toll design includes a two-part tariff where the fixed component is charged during each entry into a new zone (in other words each time a zone boundary is traversed). Overall, it results in the fixed part of the tariff being optimized at a much lower level than in the case when the zones are spatially compact (SSCOMP in Scenario **B3** and Scenarios **B4, B5**). Interestingly, the low tolls charged result in a high consumer surplus, comparable with the best performing scenario since the travel time gains and reductions in schedule delay costs are still significant.

The second reason for the poor performance of SCCEL in Scenario **B2** may be attributed to the clusters themselves. Observe that the key difference in the clusters or zones between Scenario **B2** and Scenario **B3** (which yields a significantly larger welfare) is that Scenario **B3** clearly demarcates the Back Bay region from the rest of Boston (Figure 4) whereas this is not the case in Scenario B2. The Back Bay region contains the Prudential center, which is a major attractor of trips in the morning peak and hence, arguably, the zone definitions in Scenario **B3** are more meaningful. This is also evident from the clustering performance metrics which clearly indicate that the clusters are more homogeneous in the case of Scenario **B3** than **B2** (SC of 0.198 versus 0.114).

Turning to the hierarchical density-based clustering approaches, we observe that both OPTICS (Scenario B4) and HDBSCAN* (Scenario B5) yield meaningful clusters/zone definitions. Both distinguish the densely residential South Boston region (red cluster in Figure 6a and orange cluster in Figure 6b; see also Figure 4) from the commercial South Boston Waterfront (pink cluster in Figure 6a and blue cluster in Figure 6b). In Scenario B4, the commercial downtown region (dark green cluster in Figure 6a) is separated from the more residential North End and West End regions (light green cluster in Figure 6a). These three regions are all combined into a single zone in Scenario B5 (green cluster in Figure 6b). The most notable difference in the clusters between B4 and B5 and one that most likely leads to the significant performance difference is that Scenario **B5** clearly demarcates the Back Bay region from the residential South End region (red and purple clusters in Figure 6b) unlike Scenario B4. As discussed earlier, this appears to be the reason for Scenario B5 yielding the largest gains in social welfare and consumer surplus. Scenario B5 also yields the clusters with links that are internally homogeneous (SC of 0.4). Note that when using speed for clustering as we have done, the more homogeneity within the clusters in an of itself does not appear to guarantee superior performance in terms of welfare. This is evident when comparing Scenarios B3 and B4; B4 yields superior metrics in terms of clustering performance but yields lower overall welfare. This underscores the importance of checking the reasonableness of the clusters themselves using context specific knowledge of demand patterns, land-use etc.

Notably, the TT performance improvement illustrated in Figure 7b, relative to the base case **B0** is substantial in all schemes. It would appear that the Boston CBD area would benefit from an application of a predictive distance-based tolling scheme, with average travel time TT improvements of up to 52% (relative to **B0**). However, we do caution that the large travel time improvements may also be, in part, due to a large number of short 'crossing' trips that are an artifact of modeling only the CBD area.

While it should be stated that while all the predictive distance-based tolling schemes yield substantial network performance benefits when compared to the No Toll scenario, scenario **B5** with HDBSCAN*-

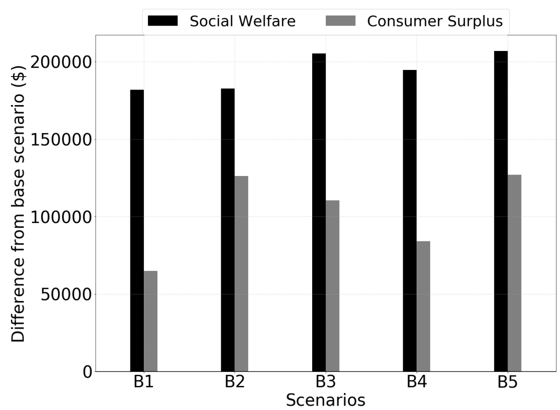
based tolling zone derivation yields the largest welfare gains. The observed welfare increase comes from the reduction in schedule delay costs and low travel times, which may be attributed to the efficient internalization of travel externality-associated costs through distance-based tolling. This in fact applies to all distance-based schemes considered in the experiments.

Scenario **B3** with only 2 tolling zones derived from SSCOMP resulted in comparable levels of performance to the Scenario **B5**, and in cases where computational effort poses a significant hurdle for practical implementation, it would be preferable to use SSCOMP. Although the HDBSCAN*-based tolling zone derivation leads to the best results, it is also more computationally intensive, due to the large number of tolling function parameters that require optimization. The overall computational time for scenarios B1-B3 were around 4 hours. Given we are simulating the 6-9 AM peak period, this does not yet achieve real-time performance for a 5-minute horizon. However, this could be attained by increasing the parallelization or marginally reducing the number of GA generations during the optimization. In case of scenarios B4, B5 where the number of zones are larger, the run times were significantly higher at 12 hours. In this case, to achieve real-time performance we would need to switch to a 15-minute roll period. For more details on computational considerations we refer the reader to Gupta et al. (2020).

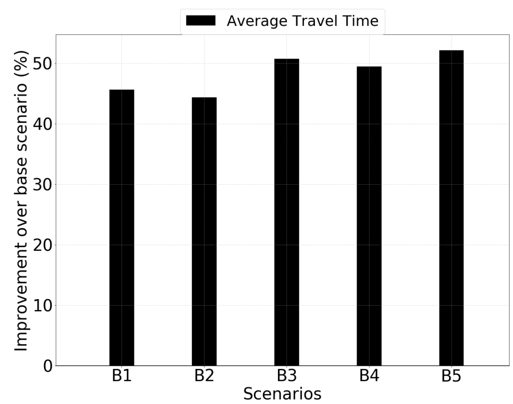
Table 4: Performance measures

	Scenarios				
Metrics	$B1$	$B2$	$B3$	$B4$	$B5$
SW ($)	181792.0	182623.5	205345.1	194744.3	206866.5
CS ($)	64814.9	126194.4	110504.2	84083.7	127062.4
TT (s)	168.0	172.2	152.3	156.3	147.9

In Figure 8a, the Empirical Cumulative Distribution Function (ECDF) of the total toll charge (for the population of vehicles) for scenarios **B1, B2, B3, B4, B5** is presented. It is evident that the majority of the traveler population (almost 90%) pay total tolls no higher than $3 for any of the distance-based pricing schemes. Further, the overall magnitude of toll charges in the case of scenario



(a) Difference in SW, CS for scenarios **B1, B2, B3, B4, B5** relative to **B0**

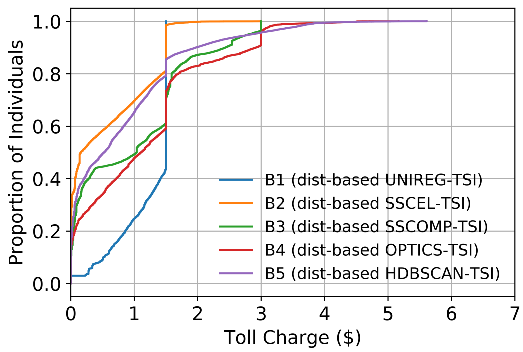


(b) Percentage Improvement in TT for scenarios **B1, B2, B3, B4, B5** relative to **B0**

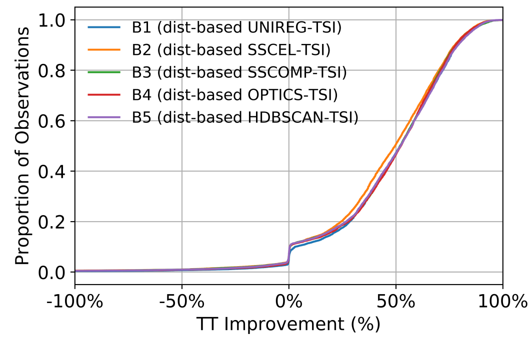
Figure 7: Performance results for scenarios **B1, B2, B3, B4, B5** relative to **B0**

B4 is consistently higher than that of scenarios **B3, B5**, which happen to be the best performing scenarios. On the other hand, the overall magnitude of toll charges in scenario **B2** is consistently lower than that of scenarios **B3, B5**, which could be explained by the fact that the corresponding tolling zone derivation suffers in terms of spatial compactness, thus leading to lower tolling efficiency. For scenario **B1**, where the entire BCBD network is treated as a single tolling zone, we can observe that more than 60% of the population is charged the toll upper bound (1.5$), which leads to inequitable charging, but also higher revenue. The reason for this lies in the fact that, unlike the case of scenarios **B2-B5**, there is only one zone for the vehicles to traverse.

As is evident from Figure 8b, demonstrating the ECDF of travel time improvement per OD-pair, for less than 10% of the traveler population, travel times are equal or lower for scenario **B0**, as compared to scenarios **B1-B5**. Up to 90% of the traveler population benefits from lower travel times, in scenarios **B1-B5** employing distance-based tolling methods, compared to base scenario **B0**. It is also evident that the largest proportion of the traveler population subset that benefits from lower travel times corresponds to **B2**, though followed very closely by **B1**.



(a) ECDF of total Toll Charge values for scenarios **B1, B2, B3, B4, B5**



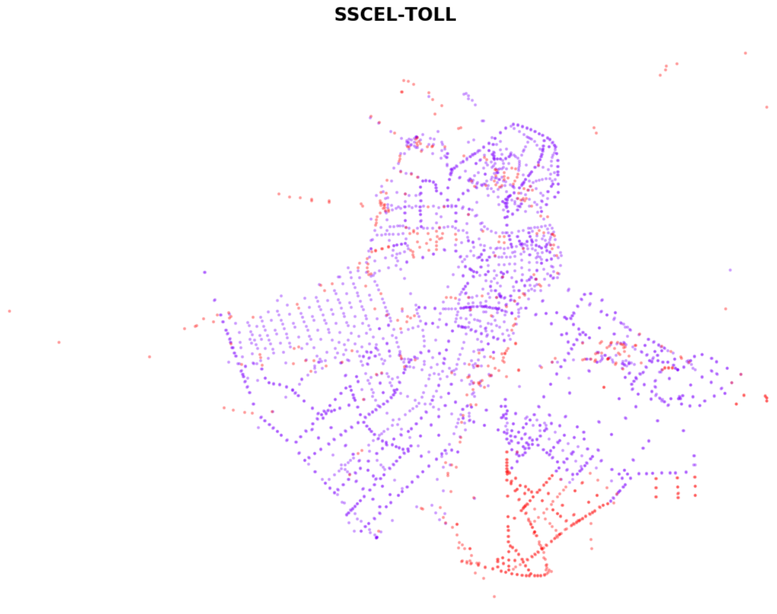
(b) ECDF of TT improvement per OD-pair over **B0** for scenarios **B1, B2, B3, B4, B5**

Figure 8: Empirical Cumulative Distribution Functions for scenarios **B1, B2, B3, B4, B5** relative to **B0**

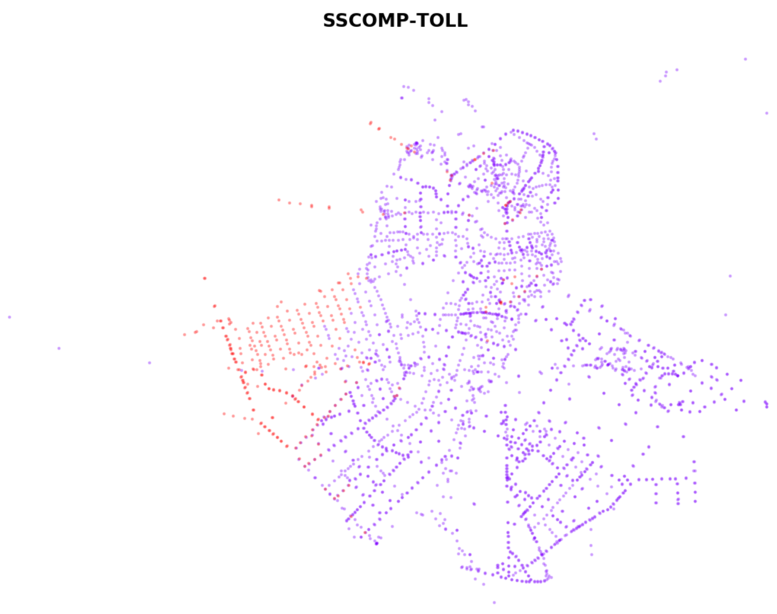
5.3. Iterating between toll zone definition and toll optimization

Recall that the motivation behind using unsupervised learning for the toll zone definition is to decouple the problems of toll definition and toll value optimization. This avoids having to solve a complex mixed-integer programming problem for the design of the distance-based scheme. Second, and more importantly, the decoupling of the two problems also serves to provide a useful separation between what is performed *offline* and what is performed *online*. Specifically, the proposed framework involves setting the tolling zones offline and then optimizing the toll values in real-time every five or fifteen minutes (within for example, a traffic management system). In this context, it may be desirable to re-evaluate the zone definitions periodically, say every month or every quarter (as is done in the current ERP system in Singapore for the setting of the toll rates). In this setting, a loop from the toll optimization and the toll design would be beneficial.

In order to do so, we redo the clustering exercise in two ways. First, we compute an implied per-distance toll rate for each link and time interval from the optimized tolling function parameters obtained via the predictive distance-based toll optimization framework. We then use the resulting toll values from each scenario as a feature and perform the clustering once again with SSCEL,SSCOMP,OPTICS and HDBSCAN*, respectively. Note that clearly, since tolling function parameters are identical for all links in the same zone by construction, this can only yield an aggregation of the original zone definitions. Nevertheless, it serves to examine robustness of the original zone definitions. Second, we redo the clustering using the travel speed indices obtained after the application of the optimized tolls. Interestingly, in both cases we observe that the original clustering results and metrics are quite robust (see Figures 9–12). However, in the case that they are not, this procedure could in principle be performed iteratively and the zone definitions could be updated.

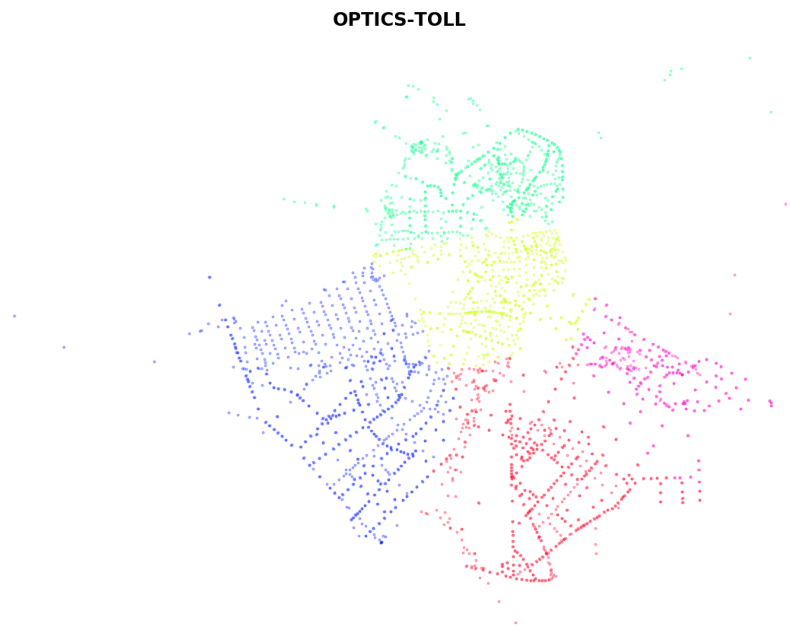


(a) 2 zones derived using SCCEL for feature TSI, with SC=0.118, DB=3.716

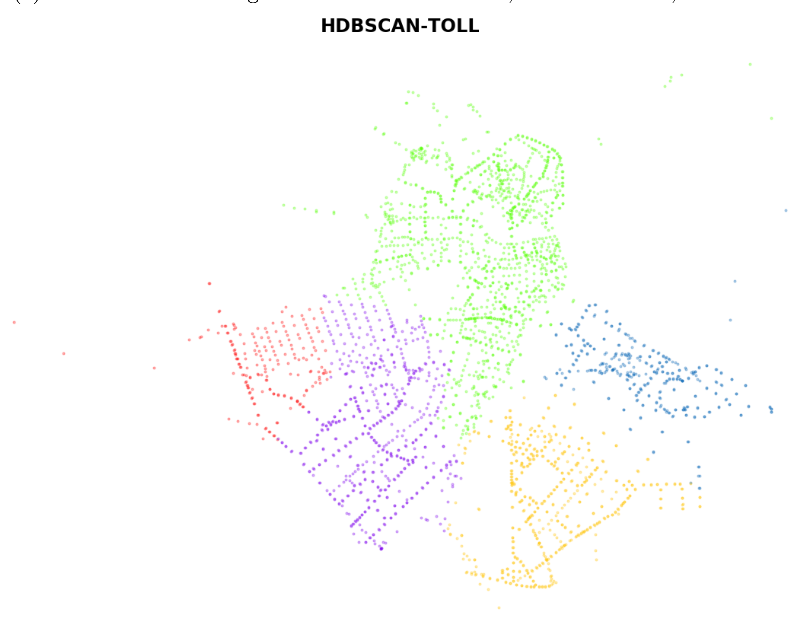


(b) 2 zones derived using SSCOMP for feature TSI, with SC=0.194, DB=1.252

Figure 9: Clustering results and tolling zones (Sparse Subspace Clustering)

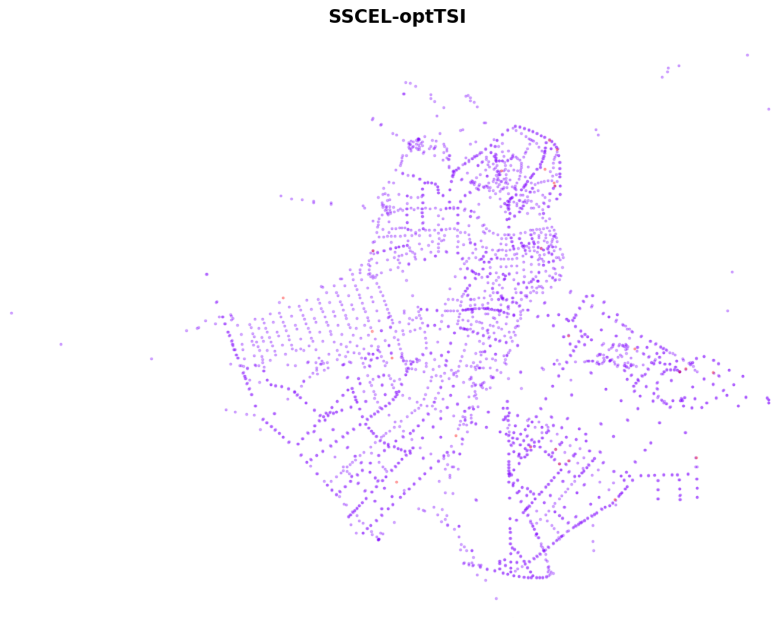


(a) 5 zones derived using OPTICS for feature TSI, with SC=0.387, DB=0.858

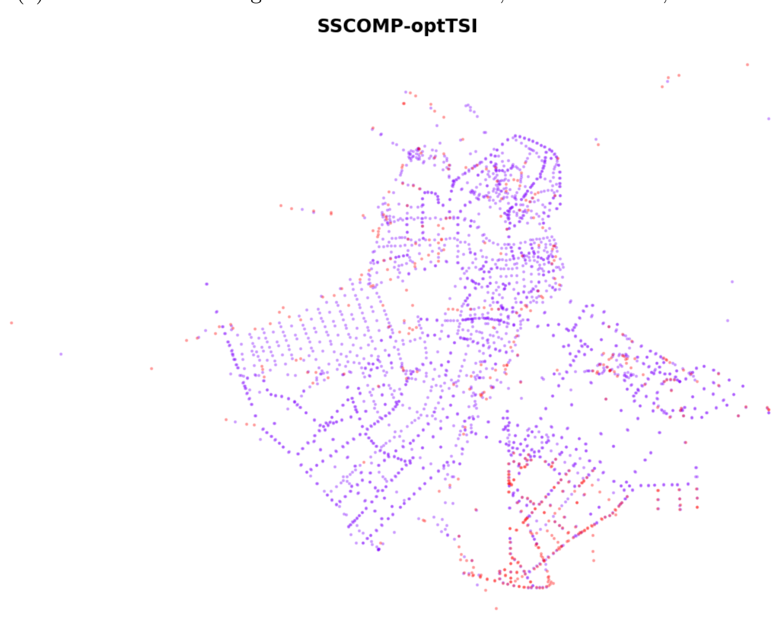


(b) 5 zones derived using HDBSCAN* for feature TSI, with SC=0.400, DB=0.776

Figure 10: Clustering results and Tolling zones (Hierarchical Density-based Clustering)

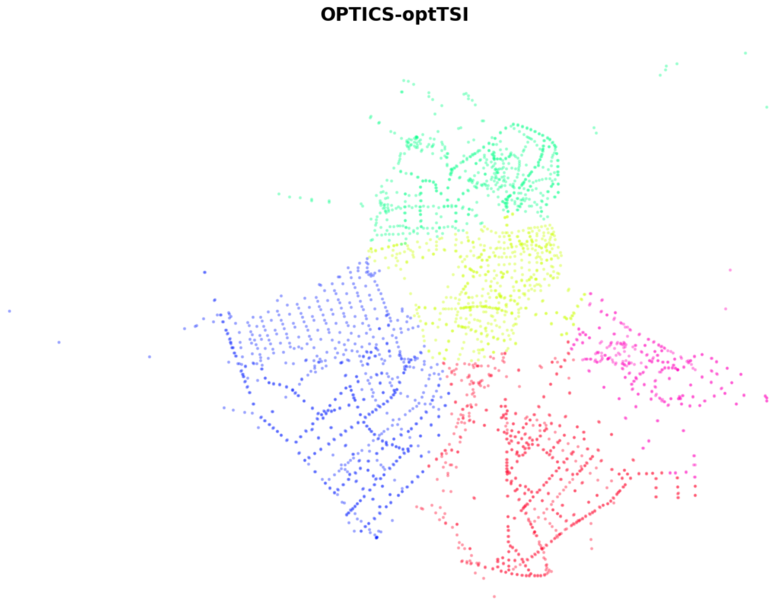


(a) 2 zones derived using SCCEL for feature TSI, with SC=0.039, DB=5.043

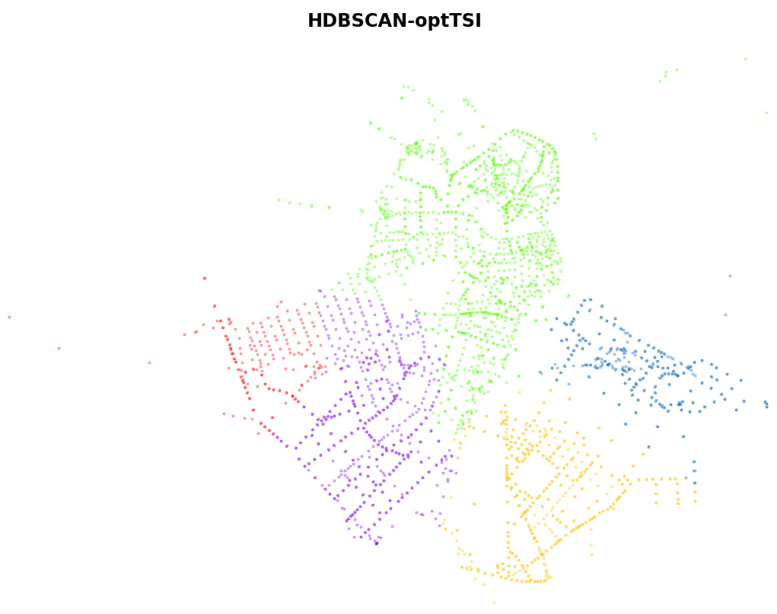


(b) 2 zones derived using SSCOMP for feature TSI, with SC=0.102, DB=4.009

Figure 11: Clustering results and tolling zones (Sparse Subspace Clustering)



(a) 5 zones derived using OPTICS for feature TSI, with SC=0.387, DB=0.858



(b) 5 zones derived using HDBSCAN* for feature TSI, with SC=0.400, DB=0.776

Figure 12: Clustering results and Tolling zones (Hierarchical Density-based Clustering)

6. Conclusions and Future Work

In this paper we investigated the use of sparse subspace clustering methods to define tolling zones for distance-based tolling schemes, and their impact on traffic network performance using a predictive real-time distance-based toll optimization framework. Experiments were conducted on the real-world urban network of the Boston Central Business District. We determined that the best network performance comes from the use of distance-based tolling zones derived from HDBSCAN*, when using Travel Speed Index data. Performance using only 2 tolling zones acquired via the SSCOMP sparse subspace clustering variant was found to be comparable to that of a 5-zone, HDBSCAN*-based derivation, so, in cases where minimizing computational effort is one of the primary objectives, as is the goal of this work, it should be considered as a viable alternative. Despite the fact that all clustering approaches produced tolling zone derivations which, as part of our framework, contributed to significant performance gains, when compared to the No Toll case, we observed large differences in performance between tolling zone derivations acquired via the sparse subspace clustering variants. Specifically, for this particular dataset, the SSCEL variant of sparse subspace clustering produced low quality clustering, due to the low degree of spatial compactness. This warrants further investigation, however, overall, tolling zone derivations acquired from both types of clustering methods, yielded significant benefits on network performance and even outperformed a predictive distance-based tolling scheme that treated the network as a single zone. Finally, the results also underscore the importance of relying not solely on clustering performance metrics but also the reasonableness of the clusters themselves using context-specific knowledge of demand patterns, land-use etc.

In future work, we aim to evaluate alternate clustering methods for systematic tolling zone derivation as part of the distance-based tolling optimization framework. Compared to our No Toll base case, social welfare and network performance results suggest that the clustering can produce distance-based tolling zones with considerable positive impact. We are also in the process of investigating alternative solution approaches, including Bayesian and Surrogate Optimization, and comparing toll optimization framework performance to that of our currently used solution approach.

Author statement

The authors confirm contribution to the paper as follows:

Antonis F. Lentzakis: Conceptualization, methodology, visualization, investigation, formal analysis, writing-Original draft preparation, writing-Review & Editing **Ravi Seshadri:** Conceptualization, methodology, writing-Original draft preparation, writing-Review & Editing **Moshe Ben-Akiva:** Conceptualization, methodology, supervision.

Acknowledgements

This research was supported by the National Research Foundation of Singapore through the Singapore-MIT Alliance for Research and Technology's FM IRG research programme.

References

Ankerst, M., Breunig, M.M., Kriegel, H.P., Sander, J., 1999. Optics: ordering points to identify the clustering structure, in: ACM Sigmod record, ACM. pp. 49–60.

Azevedo, C.L., Seshadri, R., Gao, S., Atasoy, B., Akkinepally, A.P., Christofa, E., Zhao, F., Trancik, J., Ben-Akiva, M., 2018. Tripod: sustainable travel incentives with prediction, optimization, and personalization, in: the 97th Annual Meeting of Transportation Research Board.

Bako, L., 2011. Identification of switched linear systems via sparse optimization. Automatica 47, 668–677.

Ben-Akiva, M., Koutsopoulos, H.N., Antoniou, C., Balakrishna, R., 2010. Fundamentals of Traffic Simulation. New York, NY. chapter 10-Traffic Simulation with DynaMIT. International Series in Operations Research and Management Science.

Bonsall, P.W., Palmer, I.A., 1997. Do time-based road-user charges induce risk-taking? - results from a driving simulator. Traffic Engineering and Control 38, 200–203.

Campello, R.J., Moulavi, D., Sander, J., 2013. Density-based clustering based on hierarchical density estimates, in: Pacific-Asia conference on knowledge discovery and data mining, Springer. pp. 160–172.

Chen, S., 2022. Efficient and Equitable Travel Demand Management Using Price and Quantity Controls. Ph.D. thesis. Massachusetts Institute of Technology.

Daganzo, C.F., Lehe, L.J., 2015. Distance-dependent congestion pricing for downtown zones. Transportation Research Part B 75, 91–99.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machine Intelligence 1, 224–227.

De Palma, A., Kilani, M., Lindsey, R., 2005. Congestion pricing on a road network: A study using the dynamic equilibrium simulator metropolis. Transportation Research Part A: Policy and Practice 39, 588–611.

Elhamifar, E., Vidal, R., 2009. Sparse subspace clustering, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE. pp. 2790–2797.

Elhamifar, E., Vidal, R., 2013. Sparse subspace clustering: Algorithm, theory, and applications. IEEE transactions on pattern analysis and machine intelligence 35, 2765–2781.

Eliasson, J., Mattsson, L.G., 2006. Equity effects of congestion pricing: quantitative methodology and a case study for stockholm. Transportation Research Part A: Policy and Practice 40, 602–620.

Geroliminis, N., Levinson, D.M., 2009. Cordon pricing consistent with the physics of overcrowding, in: Transportation and Traffic Theory 2009: Golden Jubilee. Springer, pp. 219–240.

Gu, Z., Saberi, M., 2019a. A bi-partitioning approach to congestion pattern recognition in a congested monocentric city. Transportation Research Part C: Emerging Technologies 109, 305–320.

Gu, Z., Saberi, M., 2019b. A simulation-based optimization framework for urban congestion pricing considering travelers' departure time rescheduling, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE. pp. 2557–2562.

Gu, Z., Shafiei, S., Liu, Z., Saberi, M., 2018. Optimal distance-and time-dependent area-based pricing with the network fundamental diagram. Transportation Research Part C: Emerging Technologies 95, 1–28.

Gupta, S., Seshadri, R., Atasoy, B., Pereira, F.C., Wang, S., Vu, V.A., Tan, G., Dong, W., Lu, Y., Antoniou, C., Ben-Akiva, M., 2016. Real time optimization of network control strategies in dynamit2. 0, in: Transportation Research Board 95th Annual Meeting.

Gupta, S., Seshadri, R., Atasoy, B., Prakash, A.A., Pereira, F., Tan, G., Ben-Akiva, M., 2020. Real-time predictive control strategy optimization. Transportation Research Record doi:10.1177/0361198120907903.

Hashemi, A., Vikalo, H., 2018. Evolutionary self-expressive models for subspace clustering. IEEE Journal of Selected Topics in Signal Processing 12, 1534–1546.

Ji, Y., Geroliminis, N., 2012. On the spatial partitioning of urban transportation networks. Transportation Research Part B: Methodological 46, 1639–1656.

Lee, H., Battle, A., Raina, R., Ng, A.Y., 2007. Efficient sparse coding algorithms, in: Advances in neural information processing systems, pp. 801–808.

Lehe, L., 2019. Downtown congestion pricing in practice. Transportation Research Part C: Emerging Technologies 100, 200–223.

Lentzakis, A.F., Seshadri, R., Akkinepally, A., Vu, V.A., Ben-Akiva, M., 2020. Hierarchical density-based clustering methods for tolling zone definition and their impact on distance-based toll optimization. Transportation Research Part C: Emerging Technologies 118, 102685.

Lentzakis, A.F., Su, R., Wen, C., 2014. Time-dependent partitioning of urban traffic network into homogeneous regions, in: Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on, IEEE. pp. 535–540.

Li, Y., Xiao, J., 2020. Traffic peak period detection using traffic index cloud maps. Physica A: Statistical Mechanics and its Applications , 124277.

Litman, T., 2019. Congestion costing critique-critical evaluation of the "urban mobility report"-9 september 2019 .

Liu, Z., Wang, S., Meng, Q., 2014. Optimal joint distance and time toll for cordon-based congestion pricing. Transportation Research Part B 69, 81–97.

LTA, 2016. Tender awarded to develop next generation electronic road pricing system.

LTA, 2021. Installation of on-board units for next-generation erp system delayed to 2023 due to global microchip shortage.

Lu, L., Xu, Y., Antoniou, C., Ben-Akiva, M., 2015a. An enhanced spsa algorithm for the calibration of dynamic traffic assignment models. Transportation Research Part C: Emerging Technologies 51, 149–166.

Lu, Y., Seshadri, R., Pereira, F., OSullivan, A., Antoniou, C., Ben-Akiva, M., 2015b. Dynamit2.0: Architecture design and preliminary results on real-time data fusion for traffic prediction and crisis management, in: Proceedings of IEEE 18th International Conference on Intelligent Transportation Systems, Spain. pp. 2250–2255.

Meng, Q., Liu, Z., Wang, S., 2012. Optimal distance tolls under congestion pricing and continuously distributed value of time. Transportation Research Part E: Logistics and Transportation Review 48, 937–957.

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

Pham, D.S., Budhaditya, S., Phung, D., Venkatesh, S., 2012. Improved subspace clustering via exploitation of spatial constraints, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE. pp. 550–557.

Rao, S., Tron, R., Vidal, R., Ma, Y., 2009. Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 1832–1845.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics 20, 53–65.

Saeedmanesh, M., Geroliminis, N., 2017. Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks. Transportation Research Procedia 23, 962–979.

Schubert, E., Sander, J., Ester, M., Kriegel, H.P., Xu, X., 2017. Dbscan revisited, revisited: why and how you should (still) use dbscan. ACM Transactions on Database Systems (TODS) 42, 19.

Simoni, M., Pel, A., Waraich, R., Hoogendoorn, S., 2015. Marginal cost congestion pricing based on the network fundamental diagram. Transportation Research Part C: Emerging Technologies 56, 221–238.

Simoni, M.D., Kockelman, K.M., Gurumurthy, K.M., Bischoff, J., 2019. Congestion pricing in a world of self-driving vehicles: An analysis of different strategies in alternative future scenarios. Transportation Research Part C: Emerging Technologies 98, 167–185.

Smith, M.J., May, A.D., Wisten, M.B., Milne, D.S., Van Vliet, D., Ghali, M.O., 1994. A comparison of the network effects of four road-user charging systems. Traffic Engineering and Control 35, 311–315.

Sun, X., Liu, Z., Thompson, R., Bie, Y., Weng, J., Chen, S., 2016. A multi-objective model for cordon-based congestion pricing schemes with nonlinear distance tolls . Journal of Central South University 23, 1273–1282.

Van Den Berg, V., Verhoef, E.T., 2011. Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. Journal of Public Economics 95, 983–992.

Yang, L., Saigal, R., Zhou, H., 2012. Distance-based dynamic pricing strategy for managed toll lanes. Transportation Research Record: Journal of the Transportation Research Board 2283, 90–99.

You, C., Li, C.G., Robinson, D.P., Vidal, R., 2016a. Oracle based active set algorithm for scalable elastic net subspace clustering, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3928–3937.

You, C., Robinson, D., Vidal, R., 2016b. Scalable sparse subspace clustering by orthogonal matching pursuit, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3918–3927.

Zhang, K., Chen, Y., Nie, Y.M., 2019. Hunting image: Taxi search strategy recognition using sparse subspace clustering. Transportation Research Part C: Emerging Technologies 109, 250–266.

Zheng, N., Rérat, G., Geroliminis, N., 2016. Time-dependent area-based pricing for multimodal systems with heterogeneous users in an agent-based environment. Transportation Research Part C: Emerging Technologies 62, 133–148.

Zheng, N., Waraich, R.A., Axhausen, K.W., Geroliminis, N., 2012. A dynamic cordon pricing scheme combining the macroscopic fundamental diagram and an agent-based traffic model. Transportation Research Part A: Policy and Practice 46, 1291–1303.

Zhu, F., Ukkusuri, S.V., 2015. A reinforcement learning approach for distance based dynamic tolling in the stochastic network environment. Journal of Advanced Transportation 49, 247–266.