Affinity Matrix Learning through Subspace Clustering for Tolling Zone Definition

Antonis F. Lentzakis
Future Urban Mobility IRG
Singapore-MIT Alliance
for Science and Technology
Singapore
antonios@smart.mit.edu

Ravi Seshadri
Future Urban Mobility IRG
Singapore-MIT Alliance
for Science and Technology
Singapore
ravi@smart.mit.edu

Moshe Ben-Akiva
Intelligent Transportation Systems Lab
Massachusetts Institute of Technology
Boston, USA
mba@mit.edu

***Abstract*—Congestion pricing is an aspect of traffic management that endeavours to alter travelers' decision-making with regards to departure time, route selection, mode choice, trip cancellation. In this paper, we propose the use of subspace clustering, as an innovative approach for defining sets of tolling zones, meant to be used in distance-based tolling implementations. We propose a new variant of Sparse Subspace Clustering by Orthogonal Matching Pursuit (SSCOMP) for learning the self-representation-based affinity matrix. Affinity Propagation clustering is applied on said affinity matrix to derive a tolling zone definition. We compare the proposed tolling zone definition approach with OPTICS, a hierarchical density-based clustering method. Travel speed indices (TSI) are used for the dataset used to perform this new variant of subspace clustering. Clustering results are used by a distance-based tolling optimization framework to evalaute network performance. Results from a Boston CBD network test case show that subspace clustering can produce tolling zone definitions with positive impact on distance-based toll optimization and overall network performance.**

***Index Terms*—sparse subspace clustering, self-representation, distance-based tolling**

I. Introduction

Urban road network partitioning is a fundamental step in traffic management, control, simulation, and policy-making. Road pricing enables traveler decision-making modification, with regards to destination selection, route selection, mode selection, departure time, trip cancellation, etc. [1]. Pricing schemes based on distance, travel-time-spent, or time-spent-in-congestion are fast becoming an attractive alternative to facility- and cordon-based pricing schemes. Distance-based pricing can address issues of safety, inflexibility and inefficiency in toll charging. Past research has investigated schemes with varying tolling function forms, linear functions are considered in [2]–[5], non-linear in [6]–[8]. The majority of distance-based toll optimization problems can be expressed as non-linear programs [4], bi-level programming models or mathematical programming models with equilibrium constraints [6], [7], and simulation-based optimization problems [5]. These have been solved using meta-heuristics [6], reinforcement learning [3], global optimization techniques [7], feeback-control approaches [5], simulation-based optimization (SBO) approaches [9]. Distance-based tolling has been investigated using nested pricing cordons [6], on idealized networks [10], using nested regions, with pricing taking place at the innermost region [5] and at the link level [11] .

Past literature for partitioning urban traffic networks used datasets based on speed, flow, density [12]–[15] and, more recently, marginal cost toll data [16]. In our case, the travel speed index (TSI) is used, a widely used quantitative indicator that employs link speed normalization [17], given the fact that identical link speed levels might reflect different traffic conditions.

Elhamifar and Vidal [18], inspired by compressed sensing [19], introduced Sparse Subspace Clustering (SSC), which makes use of the self-expressiveness property to construct the affinity matrix, upon which spectral clustering is applied, in order to derive the underlying subspaces. While subspace clustering methods have been used extensively for, among others, temporal video segmentation, switched system identification, [20], [21], only recently, has this technique come to the attention of the transportation research community. Zhang et al. [22] employed SSC to classify spatiotemporal taxi patterns with regards to their passenger searching behavior.

This paper contains the following contributions. We propose a new variant to the SSC derivative, Sparse Subspace Clustering by Orthogonal Matching Pursuit [23], hereafter termed as SSCOMP, for tolling zone definition, to be used in a distance-based tolling scheme. Travel speed indices (TSI) for each link are used to form the dataset used to derive the tolling zone definition. Rather than spectral clustering, we employ Affinity Propagation [24] in order to derive the tolling zone definition from the self-representation-based affinity matrix learned by Orthogonal Matching Pursuit. We employ a predictive distance-based toll optimization framework to assess the impact the tolling zone definition has on network performance. Experiments were conducted using the real-world Boston CBD network.

II. Tolling Zone Definition using SSCOMP and OPTICS

A. Clustering Methods

While in the majority of solution approaches for the Subspace Clustering problem, after learning the affinity matrix, spectral clustering is applied to the resulting matrix, we opted to use Affinity Propagation as our preferred clustering method. The main factors that led to this decision, are as follows:

- No prior knowledge for the number of clusters necessary
- Affinity matrix values are not required to be in a range
- Affinity matrix values can be positive or negative
- Affinity matrix can be non-symmetric

The fact that Affinity Propagation does not require the affinity matrix to be symmetric, can be especially useful in the case of urban traffic networks represented as directed graphs, where the affinity matrix can also be asymmetrical. For our proposed variant to Sparse Subspace Clustering by Orthogonal Matching Pursuit, we exploited the property of self-representation to learn the affinity matrix, to be subsequently used in our implementation of Affinity Propagation. Data self-expressiveness [25] 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:

$$\begin{aligned} &\min_{\boldsymbol{C}} \|\mathbf{C}\|_1 \\ &\textbf{s.t.} \\ &\mathbf{X} = \mathbf{XC} \\ &diag(\mathbf{C}) = 0 \end{aligned} \tag{1}$$

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

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

We can efficiently solve it using the Orthogonal Matching Pursuit algorithm, as described in [23]. 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: $\mathbf{W} = |\mathbf{C}| + \left|\mathbf{C}^T\right|$. Then, Affinity Propagation clustering is applied to $\mathbf{W}$. Affinity Propagation [24] is a clustering algorithm that finds exemplars (data centers) by passing messages between data points, in a similar manner to the Max-Product belief propagation algorithm, for a completely connected factor graph. Affinity Propagation, other than the affinity matrix $\mathbf{W}$, makes use of the responsibility message matrix $\mathbf{Y}$, that indicates the suitability of each exemplar for each data point, the availability message matrix $\mathbf{E}$, that indicates the soundness of electing an exemplar for a particular data point, based on neighboring data points' preferences. Availability can be considered as a sum of the other data points' responsibilities, but responsibilities are penalized, the further a data point is from a particular data point's neighborhood. The set of exemplar indices can then be derived from the set of matrix indices for whom the sum of availabilities and responsibilities is maximized.

For comparison to this new variant of SSCOMP, we selected the hierarchical density-based clustering method, OPTICS [26], (Ordering Points To Identify the Clustering Structure), that requires two parameters: ϵ, which describes the maximum distance radius around a particular data point and κ, describing the minimum number of data points used as a density threshold for cluster assignment. Assume $X = \{x_1, x_2, ..., x_n\}$ represents a set of data points in a metric space (X, d). We consider a data point x 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 C, with respect to ϵ and κ, is a non-empty maximal subset of X such that every pair of data points in C is density-connected. OPTICS also considers data points that are part of a more densely packed cluster, so each point is assigned a core distance that describes the distance to the κ-th nearest neighbor:

$$d_{\text{core}}^{\epsilon,\kappa}(x) = \begin{cases} \text{Undefined} & \text{if } |N_\epsilon(x)| < \kappa \\ \kappa\text{-th smallest distance to } N_\epsilon(x) & \text{otherwise} \end{cases} \tag{3}$$

The reachability-distance of data point x_i from data point x_j is either the distance between x_i and x_j, or the core distance of x_i, whichever is bigger:

$$d_{\text{reach}}^{\epsilon,\kappa}(x_i, x_j) = \begin{cases} \text{Undefined} & \text{if } |N_\epsilon(x_i)| < \kappa \\ \max(d_{\text{core}}^{\epsilon,\kappa}(x_i), d(x_i, x_j)) & \text{otherwise} \end{cases} \tag{4}$$

In our OPTICS implementation, we make use of a single global ϵ' value to extract a flat clustering.

B. *Clustering Performance Metrics*

We elected to use two standard internal evaluation indices for the clustering produced by OPTICS and SSCOMP, the Silhouette Coefficient (SC) [27] and the Davies-Bouldin index (DB) [28]. In short, SC measures two quantities, cohesion $a(x)$, which measures average distance between data points within the cluster and separation $b(x)$, which measures the minimum average distance of data points to other clusters. Then silhouette $s(x)$ is defined as:

$$s(x) = \frac{b(x) - a(x)}{max\left\{a(x), b(x)\right\}} \tag{5}$$

where $s(x) \in [-1, 1]$ takes a value -1 for incorrect clustering, around 0 for overlapping clusters and 1 for highly dense clustering. SC is the average for all data points and can be calculated using:

$$SC = \sum_{i=1}^{N} s(x_i)/N \tag{6}$$

DB is a function of the ratio of intra-cluster scatter to inter-cluster separation. The intra-cluster scatter, S_l, also known as cluster diameter, describes the average of Euclidean distances of each individual data point belonging to cluster C and the cluster centroid. If we define $d_C^{l,m}$ as the inter-cluster centroid distance, then, for number of clusters w, DB can be calculated as follows:

$$\mathrm{DB} = \frac{1}{w}\sum_{l=1}^{w} \max_{l \neq m}\left(\frac{S_l + S_m}{d_C^{l,m}}\right) \tag{7}$$

DB values closer to 0 indicate a better clustering result.

C. Clustering Results

It is imperative that our clustering methodology produces 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 our tolling optimization framework. Concretely:

- Tolling zone definitions should not exceed a certain threshold, in terms of zone set size.
- Tolling zone definitions should remain static throughout the application period
- Spatial Compactness should come about as a direct result of the high quality clustering being performed

For the reasons mentioned in Section I, we decided to use the travel speed index (TSI) for each link, calculated as follows:

$$\mathrm{TSI}_i = 1 - \frac{\nu_i}{\nu_i^f} \tag{8}$$

Where ν_i, ν_i^f the link speed and free flow speed for link i respectively. In the case of the OPTICS implementation, we used the average of TSI across five minute intervals to account for the dynamics of TSI variation. As mentioned in Section II-A, we exploited the property of self-representation to learn the affinity matrix via the SSCOMP algorithm. In the case of static partitioning schemes derived offline, this is a preferable alternative to using the average of TSI across specific intervals, given the fact that we can use the entirety of our dataset, since self-representation is amenable for use of datasets with spatiotemporal attributes [29], [30].

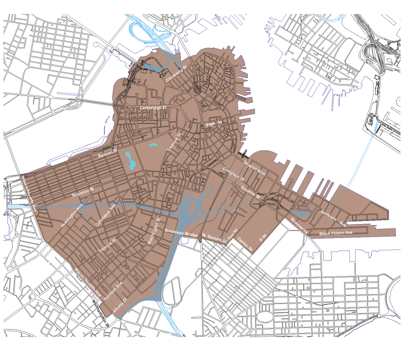
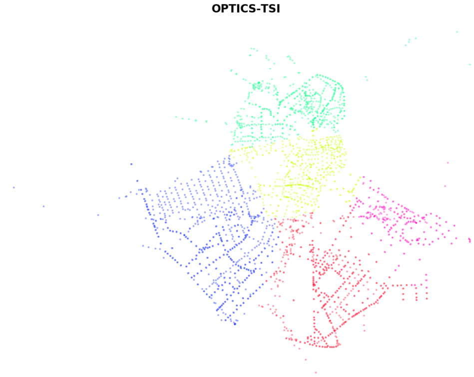
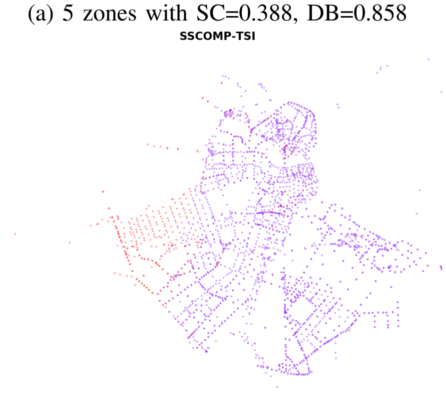


Fig. 1: Boston CBD Network

The data used for clustering include simulated speeds at the link and segment level (for the morning peak period) obtained from a calibrated DynaMIT2.0 of the Boston CBD,



(a) 5 zones with SC=0.388, DB=0.858



(b) 2 zones with SC=0.198, DB=1.276

Fig. 2: Clustering Results

shown in Figure 1. The Boston CBD network has 846 nodes, 1746 links, 3085 segments, 5057 lanes and 13080 Origin-Destination pairs. More details of the calibration process may be found in [31], [32]. After tuning the hyperparameters for OPTICS and SSCOMP, so as to produce clusterings of the Boston CBD network corresponding to tolling zone definitions which comply with our expectations for practical and feasible application, tolling zone definitions for each clustering method, as well as the clustering performance index results, are presented in Figure 2.

III. Distance-based Toll Optimization Framework

The framework for predictive distance-based toll optimization, is shown in Figure 3 using DynaMIT2.0 - a simulation-based Dynamic Traffic Assignment (DTA) system developed at the MIT Intelligent Transportation Systems Lab [33], [34]. DynaMIT2.0 comprises of the state estimation and state prediction modules. The state estimation module performs data fusion on historical database and real-time data. Current traffic network conditions are estimated through the demand/supply simulator interaction. The state prediction module predicts the network state over the prediction period and guidance information is generated, in terms of travel times, which is then passed on to the travelers during the next period. The process for strategy optimization and consistent guidance provision

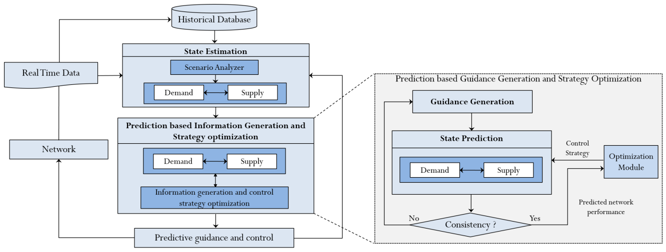


Fig. 3: Framework for Prediction-based Guidance Generation and Distance-based Strategy Optimization

has been extended in DynaMIT2.0 to optimize distance-based tolls, taking the tolling zone definition as an input. The state prediction module takes as input a candidate solution along with an initial guidance and predicts the network state through the interaction of the demand/supply simulators. Guidance is then updated, and the process continues until convergence. Subsequently, the resulting consistent guidance, along with the predicted network state is used by the optimization process to evaluate the specified objective function. Following that, the optimization process will generate a new set of control strategy solutions and the procedure will continue until optimality is achieved.

IV. Problem Description and Solution Approach

A. Tolling Function Definition

The transportation network is represented as a directed graph $G = (N, A)$, where N represents the set of n network nodes and A represents the set of m directed links. Let L be the number of tolling zones in the network G. Each tolling zone l is defined as a subset of A, i.e. $\forall l \in \{1, ..., L\}, A_l \subseteq A$, and is associated with a tolling function $\phi_l(\boldsymbol{\theta_l^t}, D_l)$ that maps the distance traveled within zone l, D_l to the toll charge, and $\boldsymbol{\theta_l^t}$ is a vector of tolling function parameters in time interval t. The toll payable within any tolling zone is assumed to be bounded, i.e $\tau_{LB} \leq \phi_l(\boldsymbol{\theta_l^t}, D_l) \leq \tau_{UB}, \forall l = 1, 2..., L \forall t = 1, 2, ..., T$.
The response of users to the predictive distance-based tolls and travel time guidance is modeled at the pre-trip level. The pre-trip choice model includes decisions of trip cancellation, mode, departure time and path and is formulated as a nested logit model (see [16] for more details).

B. Optimization and Solution

The dynamic distance-based toll optimization problem is formulated as a non-linear programming problem. The objective function for the non-linear programming problem adopted in this study is the total social welfare (SW), defined as the sum of the consumer surplus (CS) and the producer surplus (PS). CS is the sum of the experienced utilities of each traveler. PS is given by the total toll revenue (TR) minus the fixed costs (FC) and variable costs (VC). Thus, we have: $SW = CS + PS = CS + (TR - FC - VC)$.

FC is assumed to be 0, given the fact that optimal tolls are independent of the fixed costs borne by the road operator [35]. We assume that VC is a portion of the total collected toll revenue. The decision variables are the vector of tolling function parameters for the optimization period. The constraints are the DynaMIT2.0 system and upper and lower bounds on the toll values. In addition, we characterize the complex relationship of vehicle travel times and the toll function parameters and predictive guidance through a single constraint, which represents the coupled demand and supply simulators of DynaMIT2.0.

A real-coded Evolutionary Algorithm [36], a popular meta-heuristic approach amenable to parallelization, was selected for tolling function parameter optimization. This can be attributed to the highly non-convex objective function, as well as the constraints, describing the complex relationship between tolling function parameters and network travel times, having no closed-form analytical expressions.

V. Experiments

A. Design

The Boston CBD network, presented in Figure 1, will be the focus of our experiments, with historical demand and supply parameters obtained from prior DynaMIT2.0 calibrations [37]. We decided upon a bounded linear tolling function (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. Estimation interval is 5 minutes and prediction horizon is 30 minutes. Three scenarios are considered, summarized in Table I. Apart from the base *No Toll* case scenario, all scenarios involve dynamic tolls computed using the framework described in Section III. The base scenario **B0** is set up to reflect traffic conditions in the Boston CBD without any tolling scheme in place. Scenarios **B1, B2** employ predictive distance-based tolling and differ only in the way the definitions of the tolling zones were derived. In scenario **B1** (termed **OPTICS-TSI**), tolling zone definitions are derived from TSI data using OPTICS. In scenario **B2** (termed **SSCOMP-TSI**), tolling zone definitions are derived from TSI data using SSCOMP.

TABLE I: Summary of simulation scenarios

Scenario	Tolling Scheme	Description
B0	No Toll	*No tolling scheme in place*
B1	Distance-based	*Optimized predictive: OPTICS-TSI*
B2	Distance-based	*Optimized predictive: SSCOMP-TSI*

Scenarios **B0-B2** are evaluated on three performance measures, total social welfare (SW), consumer surplus (CS) and average travel time (TT) to capture the overall impact, together with the impact on travelers as individuals.

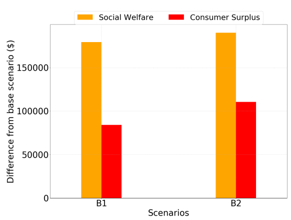
B. Results

Performance measure results are summarized in Table II. In addition, the differences in SW and CS ($) of scenarios **B1, B2**, as well as the average travel time improvement (%) relative to the base scenario B0 are illustrated in Figure 4.

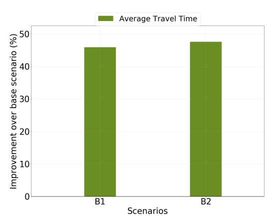
TABLE II: Performance metrics SW, CS and TT for **B1, B2**

	Scenarios	
Metrics	**B1**	**B2**
SW ($)	179285.0	189966.2
CS ($)	84083.7	110504.2
TT (s)	166.9	161.7

From Table II, **B1, B2** exhibit an increase between $179285.0 - $189966.2 and $84083.7 - $110504.2, 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 $1.8. Relative difference for scenarios **B1, B2**, in terms of SW, seem to be lower than the relative difference in terms of CS, which seems to validate the assumption that lower travel times, with subsequent lower delays, affect CS more than the total SW. What is most impressive is the TT performance improvement illustrated in Figure 4b, relative to the base case **B0**. It is evident 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 47% (relative to **B0**). While it should be stated that predictive distance-based tolling schemes with OPTICS- and SSCOMP-derived tolling zone definitions deliver substantial network performance benefits when compared to the no toll scenario, scenario **B2** with SSCOMP-derived tolling zone definition had the highest impact on distance-based tolling optimization performance.



(a) Difference in SW, CS for scenarios **B1, B2** from **B0**



(b) Improvement in TT for scenarios **B1, B2** over **B0**

Fig. 4: Performance results for scenarios **B1, B2**

VI. Conclusion

In this paper, we compared tolling zone definitions which are derived from a new variant of Sparse Subspace Clustering by Orthogonal Matching Pursuit (SSCOMP), which employs Affinity Propagation, and an established hierarchical density-based clustering method, OPTICS, with regard to their impact on adaptive distance-based toll optimization performance, when applied on an urban network setting, in our case, the Boston CBD network. Distance-based pricing scheme with tolling zones derived from SSCOMP exhibited improvements accross the gamut of performance measures (SW,CS,TT), both when compared to the base case scenario (no tolling scheme in place), as well as when compared to the performance of the distance-based pricing scheme with tolling zones derived from OPTICS. It should also be noted that the distance-based pricing scheme with SSCOMP-derived tolling zone definition manages to produce better results, even with a lower number of zones (2). In future work, we plan to investigate the possibility of integrating parking pricing to the control strategy generation process of our predictive distance-based toll optimization framework and apply subspace clustering on datasets with parking facility occupancy and road occupancy as features.

References

[1] A. de Palma and R. Lindsey, "Traffic congestion pricing methodologies and technologies," *Transportation Research Part C: Emerging Technologies*, vol. 19, pp. 1377–1399, 2011.

[2] M. A. Florian and C. D. Morosan, "A network model for capped distance-based tolls," *CIRRELT*, no. CIRRELT-2014-02, 2014.

[3] F. Zhu and S. V. Ukkusuri, "A reinforcement learning approach for distance based dynamic tolling in the stochastic network environment," *Journal of Advanced Transportation*, vol. 49, pp. 247–266, 2015.

[4] L. Yang, R. Saigal, and H. Zhou, "Distance-based dynamic pricing strategy for managed toll lanes," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2283, pp. 90–99, 2012.

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

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

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

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

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

[10] C. F. Daganzo and L. J. Lehe, "Distance-dependent congestion pricing for downtown zones," *Transportation Research Part B*, vol. 75, pp. 91–99, 2015.

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

[12] Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1639–1656, 2012.

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

[14] M. Saeedmanesh and N. Geroliminis, "Dynamic clustering and propagation of congestion in heterogeneously congested urban traffic networks," *Transportation Research Procedia*, vol. 23, pp. 962–979, 2017.

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

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

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

[18] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009, pp. 2790–2797.

[19] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *Advances in neural information processing systems*, 2007, pp. 801–808.

[20] S. Rao, R. Tron, R. Vidal, and Y. Ma, "Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1832–1845, 2009.

[21] L. Bako, "Identification of switched linear systems via sparse optimization," *Automatica*, vol. 47, no. 4, pp. 668–677, 2011.

[22] K. Zhang, Y. Chen, and Y. M. Nie, "Hunting image: Taxi search strategy recognition using sparse subspace clustering," *Transportation Research Part C: Emerging Technologies*, vol. 109, pp. 250–266, 2019.

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

[24] D. Dueck and B. J. Frey, "Non-metric affinity propagation for unsupervised image categorization," in *2007 IEEE 11th International Conference on Computer Vision*. IEEE, 2007, pp. 1–8.

[25] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2765–2781, 2013.

[26] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics: ordering points to identify the clustering structure," in *ACM Sigmod record*, vol. 28, no. 2. ACM, 1999, pp. 49–60.

[27] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[28] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 224–227, 1979.

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

[30] A. Hashemi and H. Vikalo, "Evolutionary self-expressive models for subspace clustering," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 6, pp. 1534–1546, 2018.

[31] S. Gupta, R. Seshadri, B. Atasoy, A. A. Prakash, F. Pereira, G. Tan, and M. Ben-Akiva, "Real-time predictive control strategy optimization," *Transportation Research Record*, 2020.

[32] L. Lu, Y. Xu, C. Antoniou, and M. Ben-Akiva, "An enhanced spsa algorithm for the calibration of dynamic traffic assignment models," *Transportation Research Part C: Emerging Technologies*, vol. 51, pp. 149–166, 2015.

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

[34] Y. Lu, R. Seshadri, F. Pereira, A. OSullivan, C. Antoniou, and M. Ben-Akiva, "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, 2015, pp. 2250–2255.

[35] P. Ferrari, "Road network toll pricing and social welfare," *Transportation Research Part B*, vol. 36, no. 5, pp. 471–483, 2002.

[36] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

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