Spatio-Temporal Partitioning of Large Urban Networks for Travel Time Prediction

Matej Cebecauer, Erik Jenelius, Wilco Burghout
Department of Civil and Architectural Engineering
KTH Royal Institute of Technology
Teknikringen 10, SE-100 44 Stockholm, Sweden
Email: matejc@kth.se, jenelius@kth.se, wilco@kth.se

***Abstract*—The paper explores the potential of spatio-temporal network partitioning for travel time prediction accuracy and computational costs in the context of large-scale urban road networks (including motorways/freeways, arterials and urban streets). Forecasting in this context is challenging due to the complexity, heterogeneity, noisy data, unexpected events and the size of the traffic network. The proposed spatio-temporal network partitioning methodology is versatile, and can be applied for any source of travel time data and multivariate travel time prediction method. A case study of Stockholm, Sweden considers a network exceeding 11,000 links and uses taxi probe data as the source of travel times data. To predict the travel times the Probabilistic Principal Component Analysis (PPCA) is used. Results show that the spatio-temporal network partitioning provides a more appropriate bias-variance tradeoff, and that prediction accuracy and computational costs are improved by considering the proper number of clusters towards robust large-scale travel time prediction.**

I. Introduction

Research on real-time travel time estimation and short-term prediction has until recently focused mainly on motorways and major arterials [1], [2]. Urban network travel time prediction is a more complex problem compared to motorways, due to the many additional uncertainties (e.g., signalized and unsignalized intersections, route alternatives, crossing pedestrians and flows) which are difficult to measure or predict. Further, network-wide prediction is often associated with challenges of noisy and missing data. Sophisticated models commonly applied in smaller case studies of motorways are computationally complex due to the number of inputs and parameters that have to be continuously calibrated.

GPS devices in vehicles or smart phones provide information that has proven useful for traffic management applications [3]. These opportunistic traffic sensors allow the collection of data during any time and day in urban road networks at low marginal costs [4], [3]. With the increasing availability of probe data, the literature on arterial travel time estimation and prediction has grown recently [3], [5], [6], [7]. One proposed framework considers a dynamic Bayesian network model of the spatio-temporal dependencies [8]. Another method is a spatio-temporal extension of k-nearest neighbors [9]. Online multi-output gaussian process regression is used in [10]. Network-wide travel time prediction combining advantages of Probabilistic Principal Component Analysis (PPCA) and local smoothing is proposed in [11] and extended into an integrated framework for real-time urban network travel time prediction on sparse probe data in [12].

Previous studies have shown that prediction can be improved by utilizing multivariate models over neighborhoods of links and time intervals. By pooling data from a larger neighborhood, noise can be reduced; on the other hand, variations within the neighborhood are smoothed out. Thus, there is a bias-variance trade-off where using larger neighborhoods (in the extreme, the whole network) can lower the variance, while smaller neighborhoods (in the extreme, each link individually) can lower the bias. High bias can lead to under-fitting the prediction model while high variance can lead to over-fitting.

There is a broad range of partitioning/clustering methods from different fields such as location analysis, zoning/districting, aggregation, data science, mathematics, and GIS. One stream is dealing with zoning, districting or looking for an optimal design for service systems [13], [14], [15]. A second stream considers aggregation of large problems to smaller ones in order to deal with computational complexity [16], [17]. A third stream in literature is directly connected to the transport field.

K-means [18] is one of the most popular clustering techniques thanks to its simplicity and effectiveness. Another commonly used clustering algorithm is a density-based algorithm for discovering clusters in large spatial databases with noise (DBSCAN) [19]. Recent developments in clustering for transportation networks extend normalized cut (NCut) [20] with snake similarities (S-NCut) [21]. Lopez et al. (2017) compare k-means, DBSCAN and S-NCut on a city road network and travel time data, and shows that k-means can be a good trade-off between the quality of resulting clustering and computational time [7].

The objective of this work is to investigate the effects of network partitioning on the computational time and overall accuracy of travel time prediction. To do so, a general methodology and framework combining spatio-temporal partitioning and travel time prediction is introduced. Partitioning of the links based on spatial and temporal attributes can potentially result in clusters that provide more robust and accurate prediction with reasonable bias-variance tradeoff.

The effects are studied in a case study of Stockholm, Sweden. The road network of the case study involves 11,340 heterogeneous links such as freeways/motorways, arterials and

TABLE I: Notation.

j	cluster index
J	number of clusters
$\mathcal{C}$	set of clusters
k, l	link index
K	number of links in network
i	time-of-day interval index
P	number of past time intervals to base prediction on
P_j	number of past intervals, for cluster j
I	number of time-of-day intervals
n	day index
N_E	number of days for prediction evaluation
N_C	number of days for hyper-parameter calibration
N_T	number of days for model training
d_k	link length
$\hat{v}_{ikn}$	predicted speed for link k and time interval i for day n
v_{ikn}	observed speed for link k and time interval i for day n
v_{ik}	historical average speed for link k and time interval i
lon_k, lat_k	longitude and latitude coordinates of link k
f_k	functional classification of link k
$\mathcal{G}(\mathcal{V}, \mathcal{E})$	graph of the road network
$\mathcal{V}$	set of road links (vertices)
$\mathcal{E}$	set of link connections (edges)
d_{kl}	shortest path distance between links k and l on $\mathcal{G}(\mathcal{V}, \mathcal{E})$

urban streets. To demonstrate the effects of the spatio-temporal partitioning, the data driven PPCA travel time prediction is used as the prediction method [11]. The impact of the partitioning to the motorways, major arterials and urban streets is discussed. Computational experiments reveal that partitioning can significantly improve the prediction accuracy and rapidly decrease the computational cost and time.

The paper is organized as follows. The methodology and framework combining partitioning and travel time prediction is introduced in Section II. Section III introduces the data and case study, with results reported in Section IV. Section V concludes our remarks.

II. Methodology

In this section, the framework and methodology that combine spatio-temporal partitioning and travel time prediction is presented. Subsection II-A present the concept of travel time prediction on clustered spatio-temporal data, while subsection II-B introduces the methodology behind the partitioning. Table I summarizes the notation used in this paper.

A. Travel Time Prediction and Network Partitioning

The road network is represented by the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices representing the links of the road network, and the set of edges $\mathcal{E}$ represent direct connections between links. Each vertex is characterized by the following static attributes: spatial position (lon_k, lat_k) of the midpoint of the link and functional classification f_k; and temporal attribute represented by average speed v_{ik} across the set of days in time interval i.

Let t_{ikn} be the average travel time of link k in time-of-day interval i on day n based on observations from probe data, and let d_k be the length of link k. Then $v_{ikn} = d_k/t_{ikn}$ is the mean speed of link k. This paper considers the problem of predicting $v_{i+1,kn}$ for every link k on current-day n based on observations for previous intervals across potentially all links. A general formulation that can encapsulate most proposed travel time prediction methods is

$$\hat{\mathbf{v}}_{i+1,n} = \mathbf{f}_{i,\mathcal{V}}\left(\mathbf{v}_{i-P,n}, \ldots, \mathbf{v}_{in}\right), \qquad (1)$$

where $\mathbf{v}_{in} = (v_{iln})_{l=1}^{K}$ and $P \geq 1$ is an integer parameter. Then the predicted travel time $\hat{t}_{i+1,kn}$ can be derived as $\hat{t}_{i+1,kn} = d_k/\hat{v}_{i+1,kn}$.

The motivation for network partitioning based on spatio-temporal attributes is to achieve a suitable bias-variance trade-off for multivariate travel time prediction models. Partitioning also allows for parallel computation across independent clusters. Considering J clusters $\mathcal{C} = \{C_1, C_2, \ldots C_J\}$ where $\mathcal{V} = C_1 \cup C_2 \cup \cdots \cup C_J$, the travel time predictions for every independent cluster j can be easily merged together to provide travel time prediction for all links,

$$\hat{\mathbf{v}}^j_{i+1,n} = \mathbf{f}_{ij}\left(\mathbf{v}^j_{i-P_j,n}, \ldots, \mathbf{v}^j_{in}\right), \quad j = 1, \ldots, J, \qquad (2)$$

where $\mathbf{v}^j_{in} = (v_{iln})_{l \in C_j}$ only contains observations from links belonging to cluster j. Note that different prediction models or model parameters can be applied to different clusters.

B. Partitioning Methodology

The partitioning methods used for this study may be divided into three main groups with respect to the considered attributes: *functional*, *spatial*, and *spatio-temporal* clustering.

Functional Clustering: In functional clustering each vertex k from the set $\mathcal{V}$ is assigned to a cluster based on the attribute f_k. The number of unique values across all f_k represents the resulting number of clusters J for this partitioning method.

Spatial Clustering: Two different methods for spatial clustering are studied here. In the first, each link is assigned to one of J zones that are defined before the partitioning process as spatial polygons. Links are allocated based on their spatial position. Examples of such pre-defined zones are administrative districts, municipalities, regions or other geographical or administrative units. These zones are usually historically established and they do not necessarily partition the space or road network to the most appropriate clusters.

With the second spatial clustering approach, the aim is to construct an efficient partitioning with respect to the road network distance to the centroids of the clusters. The p-median problem is one of the most frequently studied and used location problems [22], [17]. The goal is to determine exactly J centers (links in our case) in such a way that the sum of distances from all links to the closest center is minimized. All K links are considered as possible candidates for centers. The shortest road network distance between link k and l is denoted d_{kl}. Decisions are described by the set of binary variables:

$$x_{kl} = \begin{cases} 1, & \text{if link } k \text{ is assigned to center link } l \\ 0, & \text{otherwise,} \end{cases}$$

$$y_l = \begin{cases} 1, & \text{if link } l \text{ is selected as center,} \\ 0, & \text{otherwise.} \end{cases}$$

The p-median problem can be formulated as follows:

$$\text{Minimize} \qquad F=\sum_{k=1}^{K}\sum_{l=1}^{K} d_{kl}x_{kl} \tag{3}$$

subject to

$$\sum_{l=1}^{K} x_{kl}=1 \qquad k=1,2,\ldots,K \tag{4}$$

$$x_{kl}\le y_l \qquad k,l=1,2,\ldots,K \tag{5}$$

$$\sum_{l=1}^{K} y_l=J \tag{6}$$

$$x_{kl},y_l\in\{0,1\} \qquad k,l=1,2,\ldots,n. \tag{7}$$

The objective function (3) minimizes the sum of distances from all links to the centers. The constraints (4) ensure that each link is allocated to exactly one center. The constraints (5) allow allocating links only to the links selected as center, and the constraint (6) makes sure that exactly J centroids are selected. The clusters $C=\{C_1,C_2,\ldots,C_J\}$ are established with allocation of the links to the closest centers. The state-of-art algorithm [23] is used here to find the optimal solution.

Spatio-Temporal Clustering: Spatio-temporal partitioning is used to reflect the temporal effects of the link speeds v_{ik} together with spatial locations (lon_k, lat_k) in the clustering process. The well-known multidimensional clustering method k-means is used to establish clusters in this 3-dimensional space [18]. K-means is a common tool for unsupervised data classification, similarity grouping, nonlinear prediction, etc. Given a set of K observations (represented by links in our case) $(x_1,x_2,\ldots,x_K)$ in time interval i, each of which is a d-dimensional real vector, the clustering algorithm aims to partition the observations into a J-fold partitioning $C=\{C_1,C_2,\ldots,C_J\}$ which is efficient in the sense of within-class variance:

$$\arg\min_{C}\sum_{j=1}^{J}\sum_{k\in C_j}\|\mathbf{x}_k-\mathbf{y}_j\|, \tag{8}$$

where y_j represents the centroid of cluster j. Similarly to p-median, the constraint (4) ensures that each observation belong exactly to one cluster, and (6) ensures that J clusters are constructed. With k-means clustering, cluster centroids can be any point in the d-dimensional space, while for p-median it has to be one of the links.

III. Case Study

A. Network and Travel Time Data

The network partitioning and travel time prediction is applied to the road network of the Stockholm, Sweden (see Figure 1). In the digital road network each link has a number of attributes, including length, spatial position, functional road class, and speed limit. The functional class categorizes roads from class 0 (major roads) to class 9 (minor service or side streets). Functional classes up to 5 are included here; for classes higher than 6 (minor service or side streets) more than

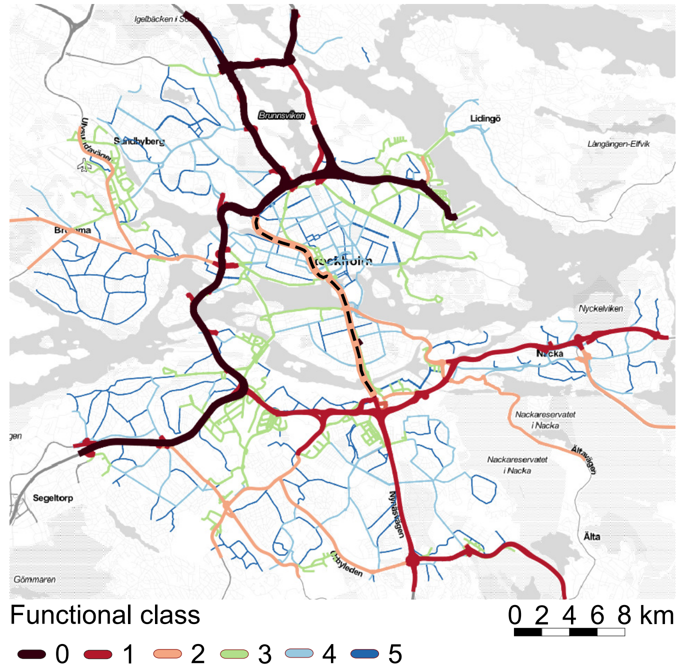

Fig. 1: Visualization of case study road network. Dashed line represents the north-south axis of Stockholm city

80% of the links observations across all days and intervals are missing. The case study road network consists of 11,340 links (functional class 0: 350 links; class 1: 764 links, links; class 2: 751 links; class 3: 2,758 links; class 4: 3,098 links and class 5: 3,619). The functional classes 0 and 1 are considered as the motorways and rest as the urban road network.

Link travel times are estimated for 15 minutes intervals using probe data from ca. 1,500 taxis operating in the area. For details on the travel time estimation method see [4], [24], although speeds (km/h) are used instead of travel times in the models. One year of data (year 2014) is used. Saturdays, Sundays and all holidays as well as school holidays are removed from the data set. After this filtering, data from $N=179$ days are used. Of these, $N_C=30$ days are randomly selected as calibration set, used for calibrating prediction model hyper-parameters. Further, $N_E=30$ randomly selected days are used for the prediction performance evaluation. The remaining $N_T=119$ days are used as training set for estimating the prediction model parameters.

B. Computational Experiments

The travel time prediction is provided for each cluster separately using the data driven PPCA method [11] which separates the prediction from the calibration of the parameters. Spatio-temporal correlations are inferred from historical data based on maximum likelihood estimation and an efficient EM algorithm for handling missing data. Prediction is performed in real-time by computing the expected distribution of link travel times in future time intervals, conditional on recent current-day observations.

TABLE II: Sets of clusters as the result of partitioning.

Set notation	Short description
no clustering	All links in a single cluster.
cluster per link	One cluster per link (11,340 clusters).
Functional clustering	
func	Each link $k = 1, \ldots, K$ is assigned to a cluster based on the f_k attribute. It results in 7 clusters (see Figure 1).
Spatial clustering	
zones_{25}	Consider 25 administrative districts of Stockholm city. Each link $k = 1, \ldots, K$ is allocated to the particular district considering its spatial position (lon_k, lat_k).
pmedian_{25}	Clusters are results of the p-median location-allocation problem for $J = 25$ centers.
pmedian_{110}	Clusters are results of the p-median location-allocation problem for $J = 110$ centers.
Spatio-functional clustering	
$\text{func\&zones}_{25}$	Combining "func" and "zones_{25}" criteria results in 110 non-empty clusters.
Spatio-temporal clustering	
kmeans_{25}	Clusters are results of k-means clustering for $J = 25$ clusters.
kmeans_{110}	Clusters are results of k-means clustering for $J = 110$ clusters.

Prediction 15 minutes (one time interval) into the future is investigated for two day-time periods: morning peak (07:45–9:00); and afternoon peak (16:45–18:00). Link directions are not considered because congestion propagates backward and traffic volume propagate forward and we would like to consider both of them like in [7].

In order to partition the set of links $\mathcal{V}$ into J clusters, where $\mathcal{V} = C_1 \cup C_2 \cup \cdots \cup C_J$, the methods presented in section II-B are applied to establish several different sets of clusters. Table II summarizes the sets used for the experiments. The set *"no clustering"* constitutes the reference set to which the results of the other methods are primarily compared. Set *"cluster per link"* represents the extreme case where each link is its own cluster, and thus is calibrated and predicted independently.

It is common in the literature to study motorways and urban road networks separately. Thus, experiments are evaluated for four different link groups: motorways (functional class 0 and 1), the Stockholm "north-south axis" (dashed links of the functional class 2 in Figure 1), major arterials (functional class 2 without dashed links), and urban streets (functional class 3,4 and 5). It is important to note that only the first group represent motorways while the other three groups represent the urban road network.

IV. Results

Prediction accuracy is evaluated in terms of the mean absolute error (MAE) between predicted $\hat{v}_{ikn}$ and observed v_{ikn} speeds in interval i across N_E days and K links:

$$\text{MAE}(i) = \frac{1}{KN_E} \sum_{k=1}^{K} \sum_{n=1}^{N_E} |\hat{v}_{ikn} - v_{ikn}| \tag{9}$$

In Figure 2 the sets *"historical mean"* and *"no clustering"* represent prediction without partitioning, to which the results

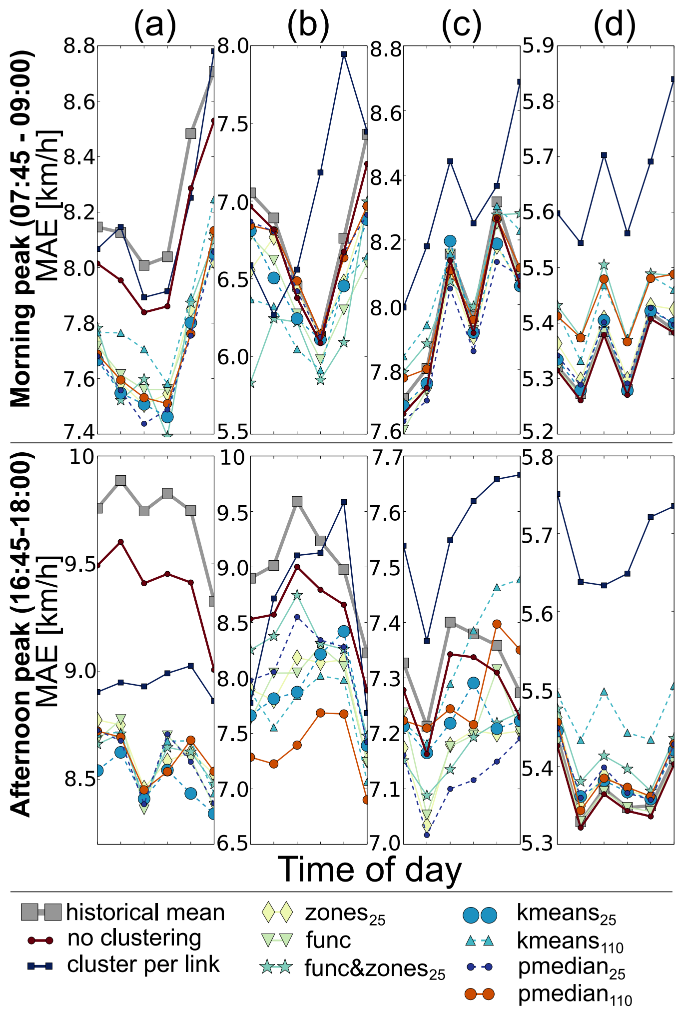

Fig. 2: Mean absolute error (MAE) of predicted vs. measured link speeds across all links of the case study. (a) motorways. (b) North-south axis. (c) Main arterial streets. (d) Lower level urban streets.

of the partitioning methods are compared. These sets include all 11,340 links of the case study together when estimating the model and calibrating the parameters. The set *"func"* produces only two link clusters for motorways and only one cluster for the north-south axis. Still, the clustering has a significant positive effect on prediction accuracy. Thus, network partitioning is helping to establish a better balance between prediction bias and variance.

The network partitioning improves the overall accuracy for motorways and major arterials (see Figure 2(a-c)). The overall improvement is significant and about 0.5 km/h for morning and 1 km/h for the afternoon peak on motorways (see Figure 2(a)), and can reach 1 km/h for the afternoon peak on the north-south axis Figure 2(b). The clustering does not improve the overall prediction accuracy for lower level urban streets (see Figure 2(d)). The *"cluster per link"* is the least robust partitioning, in particular for the urban streets

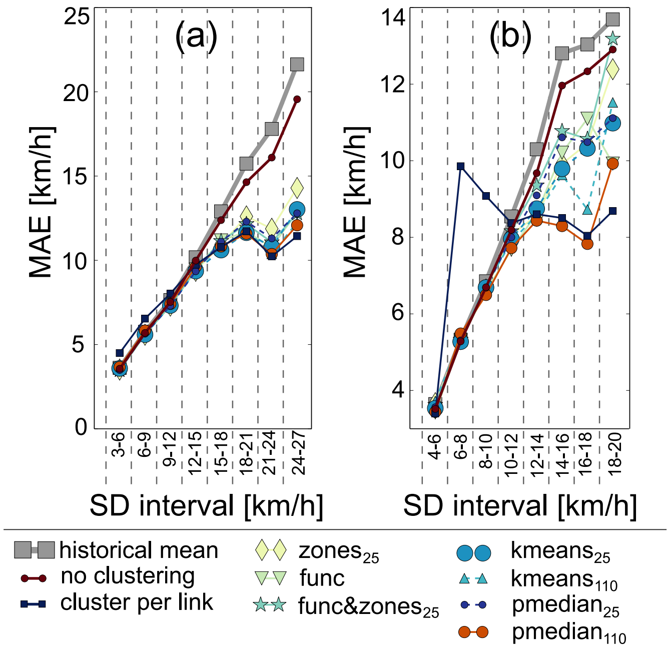

Fig. 3: Effect of the standard deviation (SD) on the prediction accuracy represented by MAE considering all intervals of the afternoon peak. Each value represent the average MAE across the links with SD (across days in N_T) in defined intervals. (a) Motorways. (b) North-south axis.

(see Figure 2(c-d)). The number of clusters has the effects on the performance. The *"no clustering"* considering whole network in one cluster can lower the variance, while *"cluster per link"* can lower the bias. Thus the appropriated number of clusters with particular network partitioning method can help to establish even more robust tradeoff between bias and variance.

It is not trivial to mark some of the clustering sets or methods as the most efficient. For example, it is clear that *"pmedian$_{110}$"* provides the most accurate travel time prediction for the north-south axis and the afternoon peak (see Figure 2(b)), while it is not very effective for the morning peak and main arterials (see Figure 2(c)). The *"pmedian$_{25}$"* is the best partitioning for the main arterials and the afternoon peak ((see Figure 2(c)).

Figure 3 reveals the relationship between the standard deviation (SD) of the links speeds across the training set and the prediction accuracy. When SD of the links grows, the mean prediction error for partitioned and non-partitioned network rises as well. Although, the increase is significantly lower for larger values of SD and all sets representing the partitioned network. The set *"cluster per link"* with links calibrated and predicted independently, has difficulties especially for links with low SD, when most likely the low variance causes the model to underfit the observed values.

Figure 4(a) shows that the spatio-temporal partitioning has significant positive effect on the computational time for model

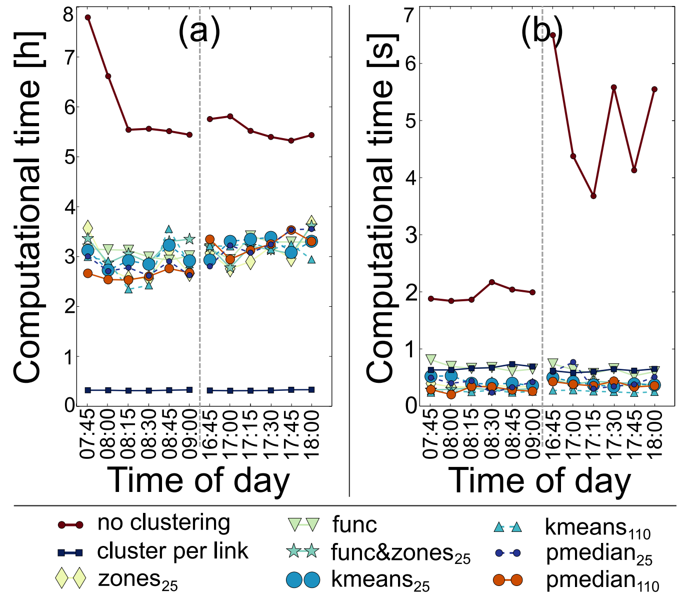

Fig. 4: Computational time for: (a) calibration of the prediction model from historical data; (b) average prediction time across evaluation days.

calibration for both peak periods, which can decrease more than 60%. The calibration time has the tendency to be lower with larger number of clusters. Further, the computational cost for prediction can be rapidly improved as shown in Figure 4(b), and can be performed in real-time (under 1 second). The improvement reaches up to 95%, when *"no clustering"* perform in 6.499 seconds and prediction for the *"kmeans$_{110}$"* set is provided in 0.275 seconds (see afternoon peak in Figure 4(b)). With the studied number of clusters, there is a tendency to lower the computational time of the prediction. The *"kmeans$_{110}$"* method is providing the most time efficient base for travel time prediction.

V. Conclusions

Travel time prediction for large urban road networks is a challenging problem due to their heterogeneous nature (e.g., signalized and unsignalized intersections, route alternatives, crossing pedestrians and flows), as well as noisy and missing data. Prediction can be improved by utilizing multivariate models over neighborhoods of links and time intervals, and there is a bias-variance trade-off where using larger neighborhoods can lower the variance but increase the bias. The present study introduces a methodology and framework for urban road network partitioning and travel time prediction. The framework is used to explore the potential and effects of spatio-temporal network partitioning on prediction accuracy and computational cost.

Several different sets of clusters are created, with respect to the introduced partitioning methods, to represent the large-scale area of the Stockholm city consisting of 11,340 links. Results reveal that network partitioning can significantly improve prediction accuracy as well as improve time efficiency and decrease the computational costs. It is not trivial to generalize which one of the presented partitioning methods is the best or the most robust, although the calibration of the most appropriate number of clusters for a particular time interval can result in an appropriate trade-off between bias and variance for overall prediction performance accuracy.

There are various future directions that can be pursued. Different clustering methods can be studied. Data fusion of different sources as probe data, loop detectors, and Bluetooth detectors can be considered for large-scale urban prediction. As one of the possible extension towards irregular incidents, more sophisticated methods based on neural networks or DTA simulation can be used for particular clusters.

Acknowledgment

The work was funded in part by the Swedish Transport Administration through the POST project, and TRENoP Strategic Research Area. The authors would like to thank Tomas Julner and Per-Olof Svensk at the Swedish Transport Administration for their support and provision of the data to the iMobility Lab at KTH.

References

[1] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.

[2] S. Oh, Y.-J. Byon, K. Jang, and H. Yeo, "Short-term travel-time prediction on highway: a review of the data-driven approach," *Transport Reviews*, vol. 35, no. 1, pp. 4–32, 2015.

[3] E. Jenelius and H. N. Koutsopoulos, "Travel time estimation for urban road networks using low frequency probe vehicle data," *Transportation Research Part B: Methodological*, vol. 53, pp. 64–81, 2013.

[4] M. Rahmani and H. N. Koutsopoulos, "Path inference from sparse floating car data for urban networks," *Transportation Research Part C: Emerging Technologies*, vol. 30, pp. 41–54, 2013.

[5] T. Hunter, A. Hofleitner, J. Reilly, W. Krichene, J. Thai, A. Kouvelas, P. Abbeel, and A. Bayen, "Arriving on time: estimating travel time distributions on large-scale road networks," *arXiv preprint arXiv:1302.6617*, 2013.

[6] Z. Zhang, Y. Wang, P. Chen, Z. He, and G. Yu, "Probe data-driven travel time forecasting for urban expressways by matching similar spatiotemporal traffic patterns," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 476–493, 2017.

[7] C. Lopez, L. Leclercq, P. Krishnakumari, N. Chiabaut, and H. Lint, "Revealing the day-to-day regularity of urban congestion patterns with 3d speed maps," *Scientific Reports*, vol. 7, no. 1, p. 14029, 2017.

[8] A. Hofleitner, R. Herring, P. Abbeel, and A. Bayen, "Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1679–1693, 2012.

[9] P. Cai, Y. Wang, G. Lu, P. Chen, C. Ding, and J. Sun, "A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 62, pp. 21–34, 2016.

[10] H. Rodriguez-Deniz, E. Jenelius, and M. Villani, "Urban network travel time prediction via online multi-output Gaussian process regression," in *Intelligent Transportation Systems (ITSC), 2017 IEEE 20th International Conference on*. IEEE, 2017, pp. 1–6.

[11] E. Jenelius and H. N. Koutsopoulos, "Urban network travel time prediction based on a probabilistic principal component analysis model of probe data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 2, pp. 436–445, 2018.

[12] M. Cebecauer, E. Jenelius, and W. Burghout, "Integrated framework for real-time urban network travel time prediction on sparse probe data," *IET Intelligent Transport Systems*, vol. 12, no. 1, pp. 66–74, 2017.

[13] H. A. Eiselt and V. Marianov, *Foundations of Location Analysis*, ser. International Series in Operations Research and Management Science. Springer, Science + Business, 2011.

[14] B. Bozkaya, E. Erkut, and G. Laporte, "A tabu search heuristic and adaptive memory procedure for political districting," *European Journal of Operational Research*, vol. 144, no. 1, pp. 12–26, 2003.

[15] K. Haase and S. Müller, "Upper and lower bounds for the sales force deployment problem with explicit contiguity constraints," *European Journal of Operational Research*, vol. 237, no. 2, pp. 677–689, 2014.

[16] R. L. Francis, T. J. Lowe, M. B. Rayco, and A. Tamir, "Aggregation error for location models: survey and analysis," *Annals of Operations Research*, vol. 167, no. 1, pp. 171–208, 2009.

[17] M. Cebecauer and L. Buzna, "A versatile adaptive aggregation framework for spatially large discrete location-allocation problems," *Computers & Industrial Engineering*, vol. 111, pp. 364–380, 2017.

[18] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.

[19] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD-96 Proceedings*, 1996, pp. 226–231.

[20] 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.

[21] M. Saeedmanesh and N. Geroliminis, "Clustering of heterogeneous networks with directional flows based on snake similarities," *Transportation Research Part B: Methodological*, vol. 91, pp. 250–269, 2016.

[22] S. L. Hakimi, "Optimum distribution of switching centers in a communication network and some related graph theoretic problems," *Operations Research*, vol. 13, no. 3, pp. pp. 462–475, 1965.

[23] S. García, M. Labbé, and A. Marín, "Solving large p-median problems with a radius formulation," *INFORMS Journal on Computing*, vol. 23, no. 4, pp. 546–556, 2011.

[24] M. Rahmani, E. Jenelius, and H. N. Koutsopoulos, "Non-parametric estimation of route travel time distributions from low-frequency floating car data," *Transportation Research Part C: Emerging Technologies*, vol. 58B, pp. 343–362, 2015.