



Machine learning and data mining for urban mobility intelligence

Etienne Come, Latifa Oukhellou, Allou Same, Lijun Sun

► To cite this version:

Etienne Come, Latifa Oukhellou, Allou Same, Lijun Sun. Machine learning and data mining for urban mobility intelligence. ESANN 2021, the 29th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Oct 2021, BRUGES, Belgium. pp 453-462, 10.14428/esann/2021.ES2021-7 . hal-03922607

HAL Id: hal-03922607

<https://hal.science/hal-03922607>

Submitted on 4 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Machine learning and data mining for urban mobility intelligence

Etienne Côme¹, Latifa Oukhellou¹, Allou Samé¹ & Lijun Sun²

1- Univ. Gustave Eiffel, COSYS, GRETTIA
F-77447 Marne-la-Vallée, France

2- McGill University,
817 Sherbrooke St. W. Montreal, QC H3A 0C3, Canada

Abstract. The last few decades have seen a faster development of digital systems for observing the mobility of people and goods. Various sensing systems - such as radio communication, Wi-Fi, Bluetooth, validation of smart cards, mobile phone, and road traffic monitoring systems - have enabled researchers and practitioners to acquire large amounts of data, which generally refer to individual and collective trajectories. The mobility data can be further enriched with side information, such as text corpora from social media, survey data, and weather information. These massive data, temporally and spatially structured, can benefit from advanced machine learning and data mining methods, providing decision aid tools, and contributing to the development of safer, cleaner, and more efficient transportation systems. They can also help to implement new mobility services for the user. This article provides an overview of methodological advances in temporal and spatial mobility data processing.

1 Introduction

This special session will be an opportunity to present recent advances in mobility data analysis. This field refers to descriptive methods, generally based on unsupervised learning, highlighting the main uses, routines, mobility patterns, and provides a basis for anomaly detection. The second field of interest in mobility data processing in the scope of this special session concerns predictive analytics, which performs in a supervised learning framework to predict quantities of interest such as transport demand or operating conditions. The implementation of these methods faces challenges related to the incompleteness, heterogeneity, and strong temporal and spatial correlation within the data, and their high dimensionality or volume.

The paper is organized as follows. Section 2 details structural time series decomposition applied to mobility data and gives an example of its application. Section 3 presents predictive models dedicated to forecast mobility demand. Section 4 focuses on tensor learning and factorization applied to mobility data.

2 Structural time series decomposition applied to mobility data

Considering mobility data as time series, structural models [1] have been used to extract from them multiple latent components, each representing an aspect of

the original profile [2, 3]. One of the advantages of such decompositions is that the resulting components can be interpreted and predicted more easily than the raw series. In particular, it becomes possible to identify specific patterns such as a long-term trend, repeating seasonal patterns (day, week, year), calendar phenomena (bank holidays ...), or the influence of exogenous phenomena (concerts, strikes, health crisis ...). In the following, we first describe these models and then show their practical use on mobility data. This work is part of the study conducted in [3].

2.1 Structural time series modeling

Given an observed time series (y_1, \dots, y_T) , with $y_t \in \mathbb{R}$, the typical structural model can be written as follows:

$$y_t = \ell_t + s_t + \sum_{j=1}^p \beta_t^{(j)} x_t^{(j)} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \quad (1)$$

where ℓ_t , s_t , ε_t are the trend, seasonal and noise components, and the $\beta_t^{(j)}$ are the coefficients associated to p known covariates $x_t^{(j)} \in \mathbb{R}$. To model their dynamic effects, the components ℓ_t , s_t and $\beta_t^{(j)}$ are usually represented in a stochastic form using auto-regressive processes [1]. Structural models are a more flexible alternative to iterative decomposition schemes such as Seasonal and Trend decomposition using Loess (STL) [4], or Census X-13ARIMA [5]. Their advantage lies in the fact that they are part of the more general framework of state-space models [6] and thus can benefit from estimation tools such as the Kalman filter, the Expectation-Maximization (EM) algorithm and Bayesian model selection criteria. However, their practical implementation requires databases collected over relatively long time periods and the incorporation of *a priori* knowledge in the model to increase their interpretive power. They have been widely used in several application fields such as economics [7], tourism [8], meteorology [9] or energy consumption [10] and have been found to be of great interest in analyzing mobility data.

The case study presented in this article concerns the decomposition of daily ridership data at a major station (La Defense Grande Arche: RER A) of the Paris railway network, which allowed us to answer questions such as:

- How do the variations in the original time series which is very often noisy translate into each component?
- What is the impact of exogenous events on passengers' decisions to use public transport?

The data were collected over a nine-year period, which constitutes an opportunity for this type of analysis. Indeed, this relatively long period of data collection allows the advent of several punctual or redundant events such as maintenance works, strikes, COVID-19 health crisis. The structural model retained for these

data has a double seasonality (weekly and annual) and has been used with some regularity constraints on the trend and seasonal component dynamics [3]. Figure 1 shows the time series and its main resulting components. As it can be seen, a time series of public transport ridership is very noisy. Its decomposition into several components to which we assign meaning is therefore useful for a better analysis of long-term and mid-term phenomena on the ridership, such as trend and seasonal variations (annual, weekly), as shown in Figure 1.

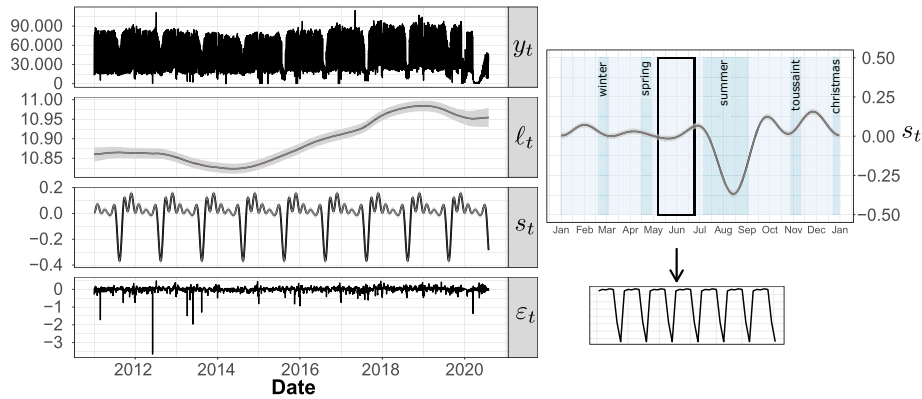


Fig. 1: Decomposition of the time series (y_t) of flows entering a railway station (La Defense Grande Arche: RER A) into log-scaled trend (ℓ_t), annual seasonality (s_t), and residuals (ε_t); Figure taken from reference [3]

The decomposition also allows a better understanding of the impact of specific exogenous events on passenger usage of public transport. Figure 2 shows for this purpose the regression component related to the Covid-19 crisis, and more particularly the lockdown and post-lockdown periods. The health crisis has strongly impacted people's movements: most of the office workers who travel daily could not go to work during this period because of the lockdown. As it can be observed, the post-lockdown period also significantly impacted the public transport usage since telecommuting and social distancing was still recommended.

It should be noticed that in mobility analyses, structural decomposition models have also been used for prediction purposes. The authors in [2] developed structural models to predict the rate of a bike sharing system usage. We can also cite the work carried out by [11], where these models were used for short-term prediction of bicycle flows. In their study on the prediction of time series of car counts at some Dublin intersections, the authors in [12] presented the decomposition of a time series into three components: trend, seasonality, and residuals. In [13], the authors used the same kind of model to forecast the monthly traffic volume of the next twelve months from a key corridor in New-Jersey.

The presented use case also opens the perspective of simultaneously modeling several time series associated with multiple stations, as in [12]. This multivari-

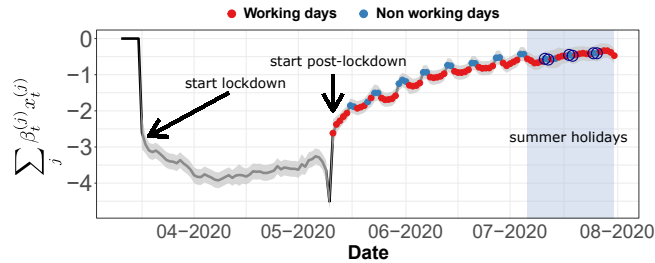


Fig. 2: Sum of regressive components associated with lockdown and post-lockdown periods and 95% confidence interval for incoming flows; Figure taken from reference [3]

ate approach could be extended to search for common dynamic factors within mobility time series.

3 Predictive models to forecast mobility demand

Predictive models and machine learnings techniques are used quite commonly to predict car traffic and have now diffused into main stream applications like google maps or waze. Such applications use floating car data at a massive scale for short term traffic forecasting. Other transportation modes also benefit from such works with other data sources. This section will detail some of the works done recently with smart card data and trains loads data in the context of public transport demand prediction and put them in perspectives with forecasting approaches used in the transportation field more generally.

Since 2004, the use of smart card data to analyse mobility in public transportation has received substantial attention from researchers. More recently, studies on mobility analysis have revolved around passenger demand forecasting. A distinction can be made between research that relates to forecasting Origin-Destination (OD) matrices and research that attempts to forecast passenger flows at a specific point or the load of a specific vehicle (e.g. train, bus). Knowledge about these factors is indeed essential for planning, operation and management in any transportation network, but each of these areas uses different types of data.

The passenger demand forecasting goals differ depending on the time horizon. Long-term forecasting aims to forecast demand with data available at the long-term period in advance (e.g., time features and planned events), which can be very useful for improving transport supply schedule. In contrast, the forecasting process can also account for the last observations, in which case it is generally referred to as short-term forecasting. In the future, in the case of an atypical situation, the main goal for transport operators is to use the forecasted passenger demand to optimize transport system operation to match transport supply to the atypical demand or propose to the passenger an alternative way to reach

their destination.

Short-term forecasting, which corresponds to a few time ahead forecasting, has been studied with different models. As an example, [14] used multiscale RBF networks to forecast the number of alighting passengers at different Beijing subway stations multiple time steps ahead ($t+15$ and $t+30$ minutes) by taking the number of boarding passengers at the other stations of the subway network into account. In this study, the authors performed an in-depth analysis of the results obtained under particular event scenarios. Other examples of subway passenger flow forecasting include the work of [15], where the authors predicted passenger flows of the next time step ($t+2$ minutes). The authors used a Bayesian network model and predicted multiple passenger flows (entry and exit) at all the stations of a subway line of the Paris network. In recent years, recurrent neural networks (RNN), in particular LSTM (introduced by [16]), have shown their efficiency in load prediction compared to standard neural networks by managing temporal dependencies [17, 18, 19, 20]. LSTM architectures are powerful in dealing with non-linear temporal dependencies in data. However, these models are usually combined with a spatial component to extract features from a network to address spatial dependencies. Thus hybrid models [19, 21, 22] divided the task into a temporal brick, most of the time an LSTM, and a spatial brick. The spatial brick can take many forms. For instance, a two-dimensional network [19] can consider spatio-temporal correlation in the traffic network. For multi-lane short-term traffic prediction, [21] replaced a standard LSTM with a convolutional LSTM to address large transport networks. Attention mechanisms (popularized by [23]) can be used in one-dimensional convolutional neural networks (1D CNN) and Gated Recurrent Units (GRU) to improve the generalization capacity of the models [22].

The above-mentioned studies are therefore built on regular time sampling, ignoring the real transport schedule. They forecast aggregated outflows and inflows over a fixed time interval for stations or metro lines. Temporal variability within the time series due to the transport schedule induces irregular temporal sampling. It makes it difficult to apply techniques that usually exploit the structural regularity of time series [18] such as previous LSTM or ARIMA frameworks. [18] used an encoder-predictor to take into account the transport schedule in their LSTM train load predictor. Another approach to deal with this specificity was introduced recently [24]. With such an approach, train load data are processed with an image-oriented methodology. First, the metro traffic is represented by generating a multi-channel image, exhibiting the spatial information of the trains running on a metro line while taking into account the irregular temporal sampling of the train loads. Input image denotes the ticket validation, time interval between trains, and waiting time at stations for all trains (in columns) and all station stops (in rows) over some time. Eventually, the output image represents the train loads. The prediction problem boils down to an inpainting / image-to-image translation problem that can be addressed with the standard deep-learning architecture used for image processing (e.g. U-net, CNN).

4 Tensor Learning and Factorization

A tensor is a multidimensional array [25], a natural way to organize urban mobility data. For example, we can represent the trip information of a metro system into a three-order (origin \times destination \times time) tensor. Each element in the tensor denotes the number of trips at a certain origin, destination, and time. Richer information, such as mode and user type, can be easily represented in a higher-order tensor. Mobility data usually exist strong repeated patterns and spatiotemporal correlation [26, 27]. Extracting these underlying patterns from high-dimensional tensors is a core step in mobility modelling, and tensor factorization is a powerful tool to achieve this goal. The idea of tensor factorization is to approximate the original tensor by the product of several much smaller matrices/tensors leveraging the low-rank property of the data. Tensor factorization can be considered as a high-order version of matrix factorization or singular value decomposition (SVD). Typical tensor factorization methods include CANDECOMP/PARAFAC (CP) decomposition [28] and tucker decomposition [29] (see [25] for a good introduction of tensor factorization). This section reviews representative works on mobility pattern mining using tensor factorization.

In a two-dimensional case, matrix factorization has been used in understanding mobility patterns, e.g., [30, 31]. In a higher-dimensional setting, an early work [32] organized the taxi trips of Beijing into a three-order (origin \times destination \times time) tensor and developed a regularized non-negative tucker decomposition method for pattern discovery. Because the factorization of a tensor is often not unique, point of interest (POI) data were used in [32] as auxiliary information to regularize the decomposition. The results showed meaningful spatial clusters and temporal structures of taxi trips. Based on different types of mobility data and regularization methods, research similar to [32] includes [33, 34]. Besides using a single matrix or tensor, [35] used a coupled decomposition to detect transit pattern, where an origin \times destination matrix and an origin \times transfer \times destination tensor share common factor loadings. The coupled approach increases the robustness of the decomposition.

The above research solves tensor factorization by minimizing the reconstruction error. On the other hand, tensor factorization can be considered from a probabilistic perspective and solved by statistical inferences [36]. For example [37] use Latent Dirichlet Allocation (LDA) to decompose the Bike Sharing System trips of Paris organized into a three-order tensor (origin \times destination \times time) and [38] applied a probabilistic tucker decomposition to understand mobility patterns of smart card transactions in Singapore. A recent work [39] integrated probabilistic tensor factorization in a Bayesian supervised learning model to predict the service choice of ride-sourcing users. The tensor factorization links the dependent variable (i.e., the type of service) with a large set of predictor variables through interpretable hidden patterns, identifying critical variables for the prediction. Besides analyzing collective mobility patterns, tensor factorization can also be used for individual-specific analysis. Similar to topic models and LDA [40], we can assume a group of people share the same

mobility patterns (in the form of factor matrices) while each person is characterized by a specific distribution over mobility patterns. Under this structure, [41, 42] applied an LDA model to extract latent topics and individual topic distributions from smart card data. The model of [41] can be used to infer unknown destinations in smart card transactions given the origin and trip time. A similar two-dimensional LDA model was applied on a license plate recognition data set for individual travel behavior pattern discovery and anomaly detection [43].

Other than mobility pattern mining, tensor factorization has a wide range of applications in more general spatiotemporal modelling tasks. Such as traffic data imputation [44, 45, 46, 47], anomaly detection [43, 48, 49], and traffic prediction [50, 51, 52, 53].

5 Conclusion

This article has detailed a set of machine learning and data mining methods to exploit mobility data. These data are rich in both spatial and temporal aspects and present some challenges for their exploitation. We have introduced three different frameworks:

- the first one, based on structural time series decomposition, is relevant when a long history of data is available and when the objective is to analyze human mobility behaviors on different time scales;
- the second one aimed at predicting travel demand in the short, medium, or long term; in this case, the issue is to develop machine learning models in a supervised framework for prediction tasks, taking into account the specificities of the domain, i.e., the particularities of the transport network, the irregular temporal sampling, or the links between the temporal series to be predicted;
- the third framework is tensor learning and factorization, which aims to extract underlying patterns from high-dimensional tensors.

The session will provide a range of current research in this area by highlighting the potential of machine learning and data mining methods to improve the movement of people in increasingly crowded cities with sustainable development criteria.

References

- [1] Andrew C Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.
- [2] Mohammed H Almannaa, Mohammed Elhenawy, and Hesham A Rakha. Dynamic linear models to predict bike availability in a bike sharing system. *International journal of sustainable transportation*, 14(3):232–242, 2020.
- [3] Paul De Nailly, Etienne Côme, Allou Samé, Latifa Oukhellou, Jacques Ferriere, and Yasmine Merad-Boudia. What can we learn from 9 years of ticketing data at a major transport hub? a structural time series decomposition. *Transportmetrica A: Transport Science*, 2021.

- [4] Robert B. Cleveland, William S. Cleveland, Jean E. McRae, and Irma Terpenning. Stl: A seasonal-trend decomposition procedure based on loess (with discussion). *Journal of Official Statistics*, 6:3–73, 1990.
- [5] Robert B. Cleveland, William S. Cleveland, Jean E. McRae, and Irma Terpenning. The x-13arima-seats seasonal adjustment program. In *Proceedings of the Federal Committee on Statistical Methodology Research Conference, Washington, DC*, 2007.
- [6] Giovanni Petris, Sonia Petrone, and Patrizia Campagnoli. Dynamic linear models. In *Dynamic Linear Models with R*, pages 31–84. Springer, 2009.
- [7] Siem Jan Koopman and Marius Ooms. Forecasting economic time series using unobserved components time series models. *The Oxford Handbook of Economic Forecasting*, 01 2012.
- [8] Jason Li Chen, Gang Li, Doris Chenguang Wu, and Shujie Shen. Forecasting seasonal tourism demand using a multiseries structural time series method. *Journal of Travel Research*, 58(1):92–103, 2019.
- [9] KV Narasimha Murthy, R Saravana, and P Rajendra. Unobserved component modeling for seasonal rainfall patterns in rayalaseema region, india 1951–2015. *Meteorology and Atmospheric Physics*, 131(5):1387–1399, 2019.
- [10] Mir Hossein Mousavi and Saleh Ghavidel. Structural time series model for energy demand in iran’s transportation sector. *Case Studies on Transport Policy*, 7(2):423–432, 2019.
- [11] Ronan Doorley, Vikram Pakrashi, Brian Caulfield, and Bidisha Ghosh. Short-term forecasting of bicycle traffic using structural time series models. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 1764–1769. IEEE, 2014.
- [12] Bidisha Ghosh, Biswajit Basu, and Margaret O’Mahony. Multivariate short-term traffic flow forecasting using time-series analysis. *IEEE transactions on intelligent transportation systems*, 10(2):246–254, 2009.
- [13] Zheyong Bian, Zhipeng Zhang, Xiang Liu, and Xiao Qin. Unobserved component model for predicting monthly traffic volume. *Journal of Transportation Engineering, Part A: Systems*, 145(12):04019052, 2019.
- [14] Yang Li, Xudong Wang, Shuo Sun, Xiaolei Ma, and Guangquan Lu. Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks. *Transportation Research Part C: Emerging Technologies*, 77:306–328, 2017.
- [15] Jérémy Roos, Stephane Bonnevey, and Gérald Gavin. Short-term urban rail passenger flow forecasting: A dynamic bayesian network approach. In *Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on*, pages 1034–1039. IEEE, 2016.
- [16] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997.
- [17] F. Toqué, E. Côme, M. K. El Mahrsi, and L. Oukhellou. Forecasting dynamic public transport origin-destination matrices with long-short term memory recurrent neural networks. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1071–1076, 2016.
- [18] Kevin Pasini, Mostepha Khouadjia, A. Samé, F. Ganansia, and L. Oukhellou. Lstm encoder-predictor for short-term train load forecasting. In *ECML/PKDD*, 2019.
- [19] Z. Zhao, W. Chen, Xingming Wu, P. Chen, and J. Liu. Lstm network: a deep learning approach for short-term traffic forecast. *Iet Intelligent Transport Systems*, 11:68–75, 2017.
- [20] Jintao Ke, Hongyu Zheng, Hai Yang, and Xiqun (Michael) Chen. Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. *Transportation Research Part C: Emerging Technologies*, 85:591–608, 2017.
- [21] Y. Ma, Z. Zhang, and A. Ihler. Multi-lane short-term traffic forecasting with convolutional lstm network. *IEEE Access*, 8:34629–34643, 2020.

- [22] Shengdong Du, Tianrui Li, Xun Gong, Zeng Yu, and Shi-Jinn Horng. A hybrid method for traffic flow forecasting using multimodal deep learning. *International Journal of Computational Intelligence Systems*, 13, 03 2018.
- [23] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [24] Thomas Bapaume, Etienne Côme, Jérémy Roos, Mostafa Ameli, and Latifa Oukhellou. Image inpainting and deep learning to forecast short-term train loads. *IEEE Access*, 9:98506–98522, 2021.
- [25] Tamara G Kolda and Brett W Bader. Tensor decompositions and applications. *SIAM review*, 51(3):455–500, 2009.
- [26] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.
- [27] Lijun Sun, Kay W Axhausen, Der-Horng Lee, and Xianfeng Huang. Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences*, 110(34):13774–13779, 2013.
- [28] J Douglas Carroll and Jih-Jie Chang. Analysis of individual differences in multidimensional scaling via an n-way generalization of “eckart-young” decomposition. *Psychometrika*, 35(3):283–319, 1970.
- [29] Ledyard R Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31(3):279–311, 1966.
- [30] Nathan Eagle and Alex Sandy Pentland. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63(7):1057–1066, 2009.
- [31] Jonathan Reades, Francesco Calabrese, and Carlo Ratti. Eigenplaces: analysing cities using the space–time structure of the mobile phone network. *Environment and Planning B: Planning and Design*, 36(5):824–836, 2009.
- [32] Jingyuan Wang, Fei Gao, Peng Cui, Chao Li, and Zhang Xiong. Discovering urban spatio-temporal structure from time-evolving traffic networks. In *Asia-Pacific Web Conference*, pages 93–104. Springer, 2014.
- [33] Yufei Han and Fabien Moutarde. Analysis of large-scale traffic dynamics in an urban transportation network using non-negative tensor factorization. *International Journal of Intelligent Transportation Systems Research*, 14(1):36–49, 2016.
- [34] Koh Takeuchi, Yoshinobu Kawahara, and Tomoharu Iwata. Structurally regularized non-negative tensor factorization for spatio-temporal pattern discoveries. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 582–598. Springer, 2017.
- [35] Bowen Du, Wenjun Zhou, Chuanren Liu, Yifeng Cui, and Hui Xiong. Transit pattern detection using tensor factorization. *INFORMS Journal on Computing*, 31(2):193–206, 2019.
- [36] Liang Xiong, Xi Chen, Tzu-Kuo Huang, Jeff Schneider, and Jaime G Carbonell. Temporal collaborative filtering with bayesian probabilistic tensor factorization. In *Proceedings of the 2010 SIAM international conference on data mining*, pages 211–222. SIAM, 2010.
- [37] Etienne Côme, Njato Andry Randriamanamihaga, Latifa Oukhellou, and Patrice Akinin. Spatio-temporal Analysis of Dynamic Origin-Destination Data Using Latent Dirichlet Allocation: Application to Vélib’ Bike Sharing System of Paris. In *TRB 93rd Annual meeting*, page 19p, France, January 2014. Transportation Research Board.
- [38] Lijun Sun and Kay Axhausen. Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transportation Research Part B Methodological*, 91:511–524, 06 2016.

- [39] Zheng Zhu, Lijun Sun, Xiqun Chen, and Hai Yang. Integrating probabilistic tensor factorization with bayesian supervised learning for dynamic ridesharing pattern analysis. *Transportation Research Part C: Emerging Technologies*, 124:102916, 2021.
- [40] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.
- [41] Zhanhong Cheng, Martin Trépanier, and Lijun Sun. Probabilistic model for destination inference and travel pattern mining from smart card data. *Transportation*, pages 1–19, 2020.
- [42] Zhan Zhao, Haris N Koutsopoulos, and Jinhua Zhao. Discovering latent activity patterns from transit smart card data: A spatiotemporal topic model. *Transportation Research Part C: Emerging Technologies*, 116:102627, 2020.
- [43] Lijun Sun, Xinyu Chen, Zhaocheng He, and Luis F Miranda-Moreno. Routine pattern discovery and anomaly detection in individual travel behavior. *Networks and Spatial Economics*, pages 1–22, 2021.
- [44] Huachun Tan, Guangdong Feng, Jianshuai Feng, Wuhong Wang, Yu-Jin Zhang, and Feng Li. A tensor-based method for missing traffic data completion. *Transportation Research Part C: Emerging Technologies*, 28:15–27, 2013.
- [45] Bin Ran, Huachun Tan, Yuankai Wu, and Peter J Jin. Tensor based missing traffic data completion with spatial-temporal correlation. *Physica A: Statistical Mechanics and its Applications*, 446:54–63, 2016.
- [46] Xinyu Chen, Zhaocheng He, Yixian Chen, Yuhuan Lu, and Jiawei Wang. Missing traffic data imputation and pattern discovery with a bayesian augmented tensor factorization model. *Transportation Research Part C: Emerging Technologies*, 104:66–77, 2019.
- [47] Xinyu Chen, Jinming Yang, and Lijun Sun. A nonconvex low-rank tensor completion model for spatiotemporal traffic data imputation. *Transportation Research Part C: Emerging Technologies*, 117:102673, 2020.
- [48] Xudong Wang, Antoine Fagette, Pascal Sartelet, and Lijun Sun. A probabilistic tensor factorization approach to detect anomalies in spatiotemporal traffic activities. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 1658–1663. IEEE, 2019.
- [49] Xudong Wang and Lijun Sun. Diagnosing spatiotemporal traffic anomalies with low-rank tensor autoregression. *IEEE Transactions on Intelligent Transportation Systems*, 2021.
- [50] Huachun Tan, Yuankai Wu, Bin Shen, Peter J Jin, and Bin Ran. Short-term traffic prediction based on dynamic tensor completion. *IEEE Transactions on Intelligent Transportation Systems*, 17(8):2123–2133, 2016.
- [51] Koh Takeuchi, Hisashi Kashima, and Naonori Ueda. Autoregressive tensor factorization for spatio-temporal predictions. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 1105–1110. IEEE, 2017.
- [52] Xinyu Chen and Lijun Sun. Bayesian temporal factorization for multidimensional time series prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [53] Yongshun Gong, Zhibin Li, Jian Zhang, Wei Liu, and Yu Zheng. Online spatio-temporal crowd flow distribution prediction for complex metro system. *IEEE Transactions on Knowledge and Data Engineering*, 2020.