Big Data-Driven Residents' Travel Mode Choice: A Research Overview

ZHAO Juanjuan[1], XU Chengzhong[2], and MENG Tianhui[1]
(1. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518000, China;
2. University of Macao, Macau SAR 999078, China)

Abstract: The research on residents' travel mode choice mainly studies how traffic flows are shared by different traffic modes, which is the prerequisite for the government to establish transportation planning and policy. Traditional methods based on survey or small data sources are difficult to accurately describe, explain and verify residents' travel mode choice behavior. Recently, thanks to upgrades of urban infrastructures, many real-time location-tracking devices become available. These devices generate massive real-time data, which provides new opportunities to analyze and explain resident travel mode choice behavior more accurately and more comprehensively. This paper surveys the current research status of big data-driven residents' travel mode choice from three aspects: residents' travel mode identification, acquisition of travel mode influencing factors, and travel mode choice model construction. Finally, the limitations of current research and directions of future research are discussed.

DOI: 10.12142/ZTECOM.201903003

http://kns.cnki.net/kcms/detail/34.1294.TN.20190919.1452.002.html, published online September 19, 2019

Manuscript received: 2019-05-09

Keywords: intelligent transportation; travel modes choice; urban computing

1 Introduction

In recent years, with the rapid economic growth and the acceleration of urbanization process in China, modern urban transportation systems, especially metropolitan transportation systems are facing a series of problems such as inadequate bearing capacity, crowded traffic and air pollution. In order to improve the efficiency of urban transport systems and promote urban sustainable development, effective management strategies have to be taken. Moreover, accurately understanding urban resident's travel mode behavior is the precondition for the government to make corresponding measures. The choice of residents' travel mode determines the distribution of people and vehicles in the urban traffic network. Since travelers are autonomous, only by sufficiently understanding travelers' choice behavior can traffic management measures be effectively formulated to lead the travelers to adjust their travel modes and further relieve traffic pressure.

This work was supported in part by National Natural Science Foundation of China (No. 61802387) and the Shenzhen Discipline Construction Project for Urban Computing and Data Intelligence.

Traditional studies on residents' travel mode choice mainly rely on field experience and sampling survey to obtain the data. However, due to limited information, it is hard to accurately describe, explain and verify residents' travel mode choice behavior. Recently, with the development of sensing technology and computing environment, long-term and continuous data can be collected, such as traffic flow, trajectory, traffic network, interest point and meteorological data. The multi-sources big data brings new perspectives to analyze and explain residents' travel mode choice behavior.

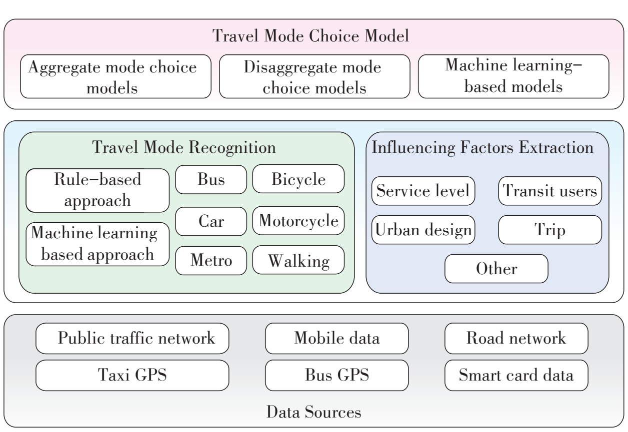
Big data-driven residents' travel mode modeling mainly contains three core issues: residents' travel mode identification, residents' travel mode influencing factors acquisition, and residents' travel mode model construction (**Fig. 1**). Among them, acquisition of resident's travel mode and influencing factors is the premise of constructing residents' travel mode model. Residents' travel mode identification identifies travelers' transportation means based on the trajectory. The traveling means include walking, bicycle, bus, subway, car, and so on. Influencing factor acquisition mainly studies the algorithm of influencing factor extraction. There are many factors affecting residents' travel mode choice behavior, such as the service level of traffic facilities, urban design, transit user, and so on. Some factors need to be extracted by fusing multiple data sources. For example, we need to integrate bus Global Positioning System (GPS), bus operation time and smart card data to extract traffic congestion, which is important for evaluating service level of traffic facilities. The resident travel mode prediction model is used to discover the relationship between these factors and travel means, so that residents' travel choice can be accurately predicted.

This paper is organized as follows. Sections 2–4 review the recent proposed methods from three aspects: residents' travel mode identification, residents' travel mode influencing factor acquisition, and residents' travel mode choice model construction. The key limitations of current research and the challenges are discussed in Section 5. Finally, we have a concluding remark in Section 6.

2 Identification of Resident Travel Mode

Residents' travel modes can be divided into personal transport modes and public transport modes. Personal transport modes include walk, bicycle, electric vehicle, car, and other modes. Public transport modes manly comprise ground bus and rail transit. Residents' travel mode identification plays an important role in understanding users' mobility and traffic situation. With the dramatic development of data processing and sensor technology (such as GPS, accelerometer, GSM and Bluetooth) over the last decade, we can collect huge amount of residents' mobility information and environmental parameters. How to use the collected data to identify urban residents' travel modes has become a hot but difficult research issue.

The process of identifying user's travel mode can be summarized as three steps. First, user's trip chain which reflects the relationship between trajectory and travel mode is extracted. Then the characteristics related to the travel mode, such as



GPS: Global Positioning System

▲ Figure 1. Framework of a big data-driven residents' travel mode choice model.

speed, start, and stop time are extracted. Finally identification methods are used to identify the user's travel mode. The identification methods can be divided into two categories: rule-based methods and machine learning-based methods.

Rule-based methods recognize different travel modes by setting different rules according to the logical characteristics of different travel modes. If the target satisfies the judgment conditions of a certain travel mode, it is classified into the travel mode. For example, STOPHER et al. [1], BOHTE et al. [2] and CHEN et al. [3] set different thresholds for average speed, maximum speed and travel time to distinguish different travel modes such as by walking, by bus and by car. However, due to its strong subjectivity (the rules and thresholds are set mainly based on experience), the accuracy of rule-based methods is not high, and the scalability is limited. So it is difficult to obtain high discriminating accuracy and scalability.

In recent years, the study of travel mode identification by machine learning, such as decision-tree, support vector machine, neural network, and stochastic forest, has become a hot research area. The accuracy and scalability have been greatly improved compared with the previous methods. The travel mode is identified using data from sensors e.g., GPS, accelerometer, Wi-Fi, and GSM. ZHENG et al. [4] collected the GPS location data of 45 objects for 6 months based on GPS device and compared the effectiveness of travel mode identification using different algorithms, including the decision-tree algorithm, Bayesian network and Support Vector Machine (SVM). The results indicated that the decision-tree algorithm is better than the other two [4]. Based on the assumption that using different travel modes make different vibrations, LAN et al. [5] used the output voltage of kinetic energy acquisition device as a signal source to detect the travel modes chosen by travelers in their daily travels. Experimental results show that this method is feasible [5]. ENDO et al. [6] put forward an automatic feature extraction method based on depth neural network method, to address the problem of low prediction accuracy caused by the artificial feature selection method and the noise of trajectory. In view of the fact that the existing machine learning methods were inadequate in explaining the results of residents' travel modes, DAS et al. [7] put forward a hybrid knowledge driven method combining fuzzy logic with neural network. A Fuzzy expert system can give a reasoning scheme but lacks the adaptability and learning ability that the neural network has. Therefore, their combination can perfectly offset mutual weaknesses [7]. Due to the ambiguous situations, for instance, traffic lights, traffic jam, bus stops and weak signal reception, current techniques report high misclassification errors for inferring transportation modes. To overcome this problem, LOPEZ et al. [8] presented a method for detecting changes of transportation mode on a multimodal journey. They used a space transformation for extracting features that identify a transition between two transportation modes based on data collected from the Google Application Programming Interface (API) for human activity classification through a crowdsourcing-based application for smartphones. The results showed improvements on precision and accuracy in comparison to initial classification data outcomes [8]. Based on multiple data sources, there are some studies on travel mode identification. FENG et al. [9] used a Bayesian network model to identify the travel model based on acceleration and GPS mixed data. The results showed that the fusion of GPS data and acceleration data has better results than that of using one single data [9]. SHIN et al. [10] proposed a real-time travel mode identification algorithm based on the acceleration and location information collected from mobile phones. JAHANGIRI et al. [11] used multiple supervised learning methods, such as k-Nearest Neighbor (kNN), SVM, decision-tree algorithm, Bagging, and Random Forest, to identify users' travel modes based on user mobile data collecting from mobile accelerometers and gyroscopes. The results revealed that Random Forest had the best prediction accuracy [11].

3 Acquisition of Influencing Factors

Residents' travel mode choice behavior is influenced by many factors. These factors can be summed up in five categories [12]–[15]: service level of transportation facilities (fare, time, comfort, and reliability); urban design (functional zones, population density, surrounding environment, public transportation density, etc.); trip (start and end location, travel purpose, and travel distance); traveler (income, car ownership, sex and age, etc.); and else (weather, special events, etc.). So far, a lot of research advances have been achieved in acquiring influencing factors based on different sensor data. In the following, we introduce the latest research progresses from three aspects.

1) In terms of service level of transportation facilities.

BARABINO et al. estimated the punctuality of buses at each stop based on bus location data and passenger's travel pattern [16], [17]. MA et al. identified and estimated the reliability of bus operation in different sections based on automatic vehicle positioning data and smart card data [18]. Based on bus location information, user's request for bus and smart card data, HADJIMITRIOU et al. identified the bus stops where the waiting time is much more than the expected by considering bus service request volume and bus time reliability [19]. CHEPURI et al. evaluated travel time variability as well as reliability using GPS based trajectory data [20]. ZHENG et al. analyzed road traffic conditions based on taxi GPS data [21]. Also based on smart card data, ZHANG et al. analyzed metro passengers' fined-grained travel time, including the time of entering and exiting metro station, transferring time, and waiting time[22]–[24]. Overcrowding is one of the major causes of discomfort. CEAPA et al. extracted station crowding patterns based on historical automated fare collection systems data [25]. WANG et al. proposed a citywide and real-time model for estimating the travel time of any path [26], based on the GPS trajectories of vehicles received in current time slots and over a period of history as well as map data sources. ZHANG et al. estimated the congestion of buses based on passengers' smart card data and bus GPS data [27]. ZHAO et al. proposed a probabilistic model to analyze how the passenger flows were dispatched to different routes and trains based on smart card and train operation timetables [28].

2) In terms of trip.

LI et al. proposed a method to predict and validate home and work locations of commuters by exploring underlying repeated travel patterns based on public transport smart card data [29]. Based on the internal spatiotemporal relationship within multi-day smart card transaction data, ZOU et al. proposed a center-point based algorithm to infer the home location for each cardholder, and a rule-based approach using the individual properties (home location and card type) of cardholders and the travel information (time and space) of each trip to identify trip purpose (work, school, shopping, and others) [30]. Furthermore, MA et al. proposed a mining method to identify transit commuters by leveraging spatial clustering and multi-criteria decision analysis approaches based on transit smart card data [31]. HUANG et al. presented an approach using spatial temporal attractiveness of Point of Interests (POIs) to identify activity-locations as well as durations from raw GPS trajectory [32]. FURLETTI et al. proposed a probability model to identify traveler's purposes, such as going to work and studying, based on GPS trajectory and interest point data [33]. LIU et al. provided a practical framework for inferring the trip purposes of taxi passengers, where the probability of points of interest to be visited is modeled by Bayes' rules, with both spatial and temporal constraints taken into consideration [34].

3) In terms of urban design.

LIU et al. proposed a scene classification framework to identify dominant urban land use type at the level of traffic analysis zone by integrating probabilistic topic models and support vector machine [35]. YUAN et al. identified urban functional zones based on the data of residents' movement trajectory, POI, urban road network and so on [36]. YAO et al. established a novel framework to map urban population distributions at the building scale by integrating multisource geospatial big data [37].

4 Residents' Travel Mode Choice Model

Constructing residents' travel mode selection model is a classical research topic in traffic planning [38]. According to the analysis unit that is an individual or a group, there are two types of models, the aggregate model and disaggregate model. The aggregate model is proposed to model the behavior of a group of travelers. The disaggregate model is put forward later based on the utility theory and its analysis unit is an individual. Compared with the aggregate model, the disaggregate model has the characteristics of high prediction accuracy, which makes the disaggregate model the main research direction in recent years.

The disaggregate model offers substantial advantage over the aggregate model as it models the behavior of individuals. The Logit model based on the theory of utility maximization is the first proposed disaggregate model. It is assumed that a traveler preference for a certain travel mode can be described by the "utility value" of the travel mode characteristics and socio-demographic attributes of the traveler. In the Logit model, the utility of an individual selecting a travel mode j is denoted as U_j and composed of two parts: the deterministic utility V_j of observable factors and the random utility ξ_j of unobservable factors. The random term ξ_j is assumed to be independent and subject to the same probability distribution (Gumbel distribution or extreme distribution). The Logit model is widely used to analyze the distribution of transport vehicles for inter-city and inner-city travelers [39], [40].

Because the Logit model assumes that the unobservable random utility ξ_j of choice branches are independent and subject to Gumbel distribution, there is no correlation among all the choices (Independence from Irrelevant Alternative). However, in some cases it is divorced from reality [41], [42]. In order to overcome the defects of the Logit model, many researchers have tried to improve the disaggregate model and proposed various and more advanced models. MINAL et al have given a survey of the research advances [43] about these models, such as Probit model and advanced Logit model. The Probit model assumes that ξ_j are not independent with each other, but this causes more parameters to calculate. When choice branches are greater than 2, it is very complex to solve the parameters. Although many approximation solutions, such as the simulation methods (Monte-Carlo simulation and McFadden), are proposed later, the qualitative analysis has uncertainty, inestimable error, and defects of complex parameter calibration process, which hinders the practicability of the model [44]–[46]. Therefore, some scholars have turned to study improved Logit models which have partial advantages of the Probit model. These improved Logit models can be generalized into two categories, the hierarchical Logit improvement model and the direct Logit improvement model. The nested Logit model is based on the multi-level partition of a travel mode, which is closer to reality than the multiple Logit model because of the correlation between the modes [43], [47]. The Dogit model is the representative of direct Logit models [48], which divides residents' choice of travel modes into two types, the forced choice behavior and free choice behavior. It assumes that the ratio of the forced choice behavior is fixed, which the free choice behavior is subject to the Logit model.

Along with the development of machine learning, some researchers use artificial neural networks (ANNs) to predict residents' travel mode choice [49], [50]. Compared with the Logit model, ANN models are easily applicable with their higher capability to identify nonlinear relationships between inputs and designated outputs to predict choice behaviors. OMRANI et al. [51] presented four machine learning methods, namely artificial neural net-MLP, artificial neural net-RBF, multinomial logistic regression, and support vector machines, for predicting travel mode of individuals. The results reveal that the artificial neural networks perform better compared to other alternatives [51]. Moreover, LEE et al. [52] investigated the capability of four ANN models, including backpropagation neural networks (BPNNs), radial basis function networks (RBFNs), probabilistic neural networks (PNNs), and clustered probabilistic neural networks (CPNNs), and compared their prediction performance with a conventional multinomial logit model (MNL) for mode choice problems. The results show that ANN models outperform MNL, with prediction accuracies around 80% compared with 70% for MNL [52]. ANNs perform better than tradition models on modeling residents' travel mode choice, but they have the shortage that the parameters are hard to explain, which blocks its popularization in transportation applications.

5 Limitations of Current Research and Challenges

By now, a large number of related studies have been proposed to model residents' travel mode choice behavior. However, due to incomplete data, complex influence factors, unclear relationship between factors and residents' travel modes, there are still some limitations in current research stage, which we summarized as the following three points:

1) Residents' travel mode identification.

Current methods are able to achieve higher accuracy in identifying the travel mode based on relatively high-frequency sampling data, but they are not applicable to identify the travel modes of all residents in a whole city. First, in general, it is difficult to collect high-frequency location data of a large number of residents. Though we can obtain some time and space information of large number residents through various sensors, such as mobile phones and smart cards, for example, phone users do not use mobile phones all the time. The smart card can only capture a user's time and space information when he/she is using public transportation. Furthermore there is no unique identification between different sensors for the same user. That is, the method cannot be directly used when there is no user association between sensors and discontinuous sparse location data. Overall, current methods cannot be directly used under the condition of sparse and discontinuous location data, and no unique identification between different but related sensors.

2) Acquisition of influencing factors.

In terms of influencing factor acquisition, there are limitations in the following two aspects: a) Due to the sparsity and incompleteness of data, it is difficult to assess the influencing factors with great accuracy based on a single source or a small number of data sources. For example, the train capacity rate and passenger flow distribution in a metro network are the important indicators in comfort evaluation. The premise is to capture passengers' route choice behavior in the subway network. However, we can only obtain each passenger's time and station information when he/she enters and leaves metro stations. If such data as cell phone signaling of a passenger is integrated, more fine-grained location information of the passenger can be obtained, thus making a more accurate estimation of route choice behavior. ZHENG et al. have made a comprehensive overview of cross-domain data fusion technology, but relatively less in the analysis of residents' travel modes [36]. b) The analysis results are dispersed and do not form a whole area. There are correlations between these factors, but current studies only focus on analyzing one or some of them, instead of paying more attention to the correlations between different factors.

3) Construction of residents' travel mode choice model.

Many researchers are studying the construction of travel mode choice models, but there have been few accurate estimations of the residents' travel modes due to limited data as well as the large gap between assumption and reality. For example, the most common definition of "utility value" is to consider the combination of fare/income, time in the car and walking time, while ignoring the influence of other factors.

6 Conclusions and Future Research

The emergence of big data provides us the possibility to deeply understand residents' travel mode choice behavior. This paper summarizes the current research status from three aspects: residents' travel mode identification, influencing factors acquisition and residents' travel mode choice model construction. The future research directions include:

1) Knowledge fusion. The methods for residents' travel mode identification and influencing factor acquisition based on a single data have been well explored. However, the methodology that can learn mutually reinforced knowledge from multiple data sources is still missing. Therefore, how to integrate heterogeneous cross-domain data sources to identify residents' travel modes and extract potential influencing factors is a challenge that needs to be addressed in the future.

2) Residents' travel mode prediction. Residents' travel mode choice behavior is affected by multiple complex factors, including internal and external environments. The correlations between these factors, as well as between factors and resident travel modes, are still unclear. Therefore, how to discover these correlations and make an accurate prediction for residents' travel mode choice is another challenge that needs to be studied in the future.

References

[1] STOPHER P, FITZGERALD C, ZHANG J. Search for a Global Positioning System Device to Measure Person Travel [J]. Transportation Research Part C Emerging Technologies, 2008, 16(3): 350–369. DOI: 10.1016/j.trc.2007.10.002

[2] BOHTE W, MAAT K. Deriving and Validating Trip Purposes and Travel Modes for Multi-Day GPS-Based Travel Surveys: A Large-Scale Application in the Netherlands [J]. Transportation Research Part C, 2009, 17(3): 285–297. DOI: 10.1016/j.trc.2008.11.004

[3] CHEN C, GONG H, LAWSON C, et al. Evaluating the Feasibility of a Passive Travel Survey Collection in a Complex Urban Environment: Lessons Learned from the New York City Case Study [J]. Transportation Research Part A Policy & Practice, 2010, 44(10): 830–840. DOI: 10.1016/j.tra.2010.08.004

[4] ZHENG Y, LIU L, WANG L, et al. Learning Transportation Mode from Raw GPS Data for Geographic Applications on the Web [C]//International Conference on World Wide Web, WWW 2008. Beijing, China, 2008:247–256. DOI: 10.1145/1367497.1367532

[5] LAN G, XU W, KHALIFA S, et al. Transportation Mode Detection Using Kinetic Energy Harvesting Wearables [C]//IEEE International Conference on Pervasive Computing and Communication Workshops. Sydney, Australia, 2016: 1–4. DOI: 10.1109/PERCOMW.2016.7457048

[6] ENDO Y, TODA H, NISHIDA K, et al. Deep Feature Extraction from Trajectories for Transportation Mode Estimation [M]. Advances in Knowledge Discovery and Data Mining. New York, USA: Springer International Publishing, 2016. DOI: 10.1007/978-3-319-31750-2_5

[7] DAS R D, WINTER S. Detecting Urban Transport Modes Using a Hybrid Knowledge Driven Framework from GPS Trajectory [J]. International Journal of Geo-Information, 2016, 5(11): 207. DOI: 10.3390/ijgi5110207

[8] LOPEZ A J, OCHOA D, GAUTAMA S. Detecting Changes of Transportation-Mode by Using Classification Data [C]//IEEE International Conference on Information Fusion. San Diego, USA, 2015: 2078–2083

[9] FENG T, TIMMERMANS H J P. Transportation Mode Recognition Using GPS and Accelerometer Data [J]. Transportation Research Part C Emerging Technologies, 2013, 37(3): 118–130. DOI: 10.1016/j.trc.2013.09.014

[10] SHIN D, ALIAGA D, TUNÇER B, et al. Urban Sensing: Using Smartphones for Transportation Mode Classification [J]. Computers Environment & Urban Systems, 2015, 53: 76–86

[11] JAHANGIRI A, RAKHA H A. Applying Machine Learning Techniques to Transportation Mode Recognition Using Mobile Phone Sensor Data [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5): 2406–2417. DOI: 10.1109/TITS.2015.2405759

[12] ZHAO F, LI M T, CHOW L F, et al. FSUTMS Mode Choice Modeling: Factors Affecting Transit Use and Access [R]. Tampa, USA: University of South Florida, 2002

[13] RACCA D P, RATLEDGE E C. Factors that Affect and/or can Alter Mode Choice [R]. Newark, USA: University of Delaware, 2004.

[14] YE X, PENDYALA R M, GOTTARDI G. An Exploration of the Relationship Between Mode Choice and Complexity of Trip Chaining Patterns [J]. Transportation Research Part B Methodological, 2007, 41(1): 96–113. DOI: 10.1016/j.trb.2006.03.004

[15] MADHUWANTHI M, MARASINGHE A, RAJAPAKSE R P C J, et al. Factors Influencing to Travel Behavior on Transport Mode Choice—A Case of Colombo Metropolitan Area in Sri Lanka [J]. International Journal of Affective Engineering, 2015, 15(2): 63–72. DOI: 10.5057/ijae.IJAE-D-15-00044

[16] BARABINO B, FRANCESCO M D, MOZZONI S. Rethinking Bus Punctuality by Integrating Automatic Vehicle Location Data and Passenger Patterns [J]. Transportation Research Part A, 2015, 75: 84–95. DOI: 10.1016/j.tra.2015.03.012

[17] BARABINO B, FRANCESCO M D, MURRU R. An Offline Framework for Reliability Diagnosis by Automatic Vehicle Location Data [C]//Conference on Advanced Systems in Public Transport. Rotterdam, Netherlands, 2015. DOI: 10.1109/tits.2016.2581024

[18] MA Z, FERREIRA L, MESBAH M, et al. Modelling Bus Travel Time Reliability Using Supply and Demand Data from AVL and Smart Card Systems [J]. Transportation Research Record Journal of the Transportation Research Board, 2015, 2533:17–27

[19] HADJIDIMITRIOU S N, KAPARIAS I, DELL'AMICO M. Investigating Urban Bus Travel Time Reliability Patterns in London Using iBus Automatic Vehicle Locating and Live Bus Arrivals data [C]//6th Symposium of the European Association for Research in Transportation. Haifa, Israel, 2017

[20] CHEPURI A, JAIRAM R, ARKATKAR S, et al. Travel Time Reliability-Based Performance Indicators Assessment for Bus Routes Using GPS-Based Bus Trajectory under Mixed Traffic Conditions [C]//3rd Conference of Transportation Systems Engineering and Management (CTSEM). Bangalore, India, 2016

[21] ZHENG Y, LIU Y, YUAN J, et al. Urban Computing with Taxicabs [C]//13th ACM International Conference on Ubiquitous Computing. Beijing, China, 2011: 89–98. DOI: 10.1145/2030112.2030126

[22] WAITING E, HOWEVER A S F. Splitting Travel Time Based on AFC Data: Estimating Walking, Waiting, Transfer, and In-Vehicle Travel Times in Metro System [J]. Discrete Dynamics in Nature and Society, 2015(1): 1–11. DOI:

10.1155/2015/539756
[23] ZHANG F, ZHAO J, TIAN C, et al. Spatiotemporal Segmentation of Metro Trips Using Smart Card Data [J]. IEEE Transactions on Vehicular Technology, 2016, 65(3): 1137–1149. 10.11 DOI: 09/TVT.2015.2409815
[24] YUE Z, CHEN F, WANG Z, et al. Classifications of Metro Stations by Clustering Smart Card Data Using the Gaussian Mixture Model [J]. Urban Rapid Rail Transit, 2017, 30(2): 48–51
[25] CEAPA I, SMITH C, CAPRA L. Avoiding the Crowds: Understanding Tube Station Congestion Patterns from Trip Data [C]//ACM SIGKDD International Workshop on Urban Computing. Beijing, China, 2012: 134–141. DOI: 10.1145 / 2346496.2346518
[26] WANG Y L, ZHENG Y, XUE Y X. Travel Time Estimation of a Path Using Sparse Trajectories [C]//Knowledge Discovery and Data Mining (KDD' 14). New York, USA, 2014: 25–34. DOI: /10.1145/2623330.2623656
[27] ZHANG J, YU X, TIAN C, et al. Analyzing Passenger Density for Public Bus: Inference of Crowdedness and Evaluation of Scheduling Choices [C]//IEEE 17th International Conference on Intelligent Transportation Systems. Qingdao, China, 2014: 2015–2022. DOI: 10.1109/ITSC.2014.6958000
[28] ZHAO J, ZHANG F, TU L, et al. Estimation of Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems [J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(4): 790–801. DOI: 10.1109 / TITS.2016.2587864
[29] LI G, YU L, NG W S, et al. Predicting Home and Work Locations Using Public Transport Smart Card Data by Spectral Analysis [C]//IEEE 18th International Conference on Intelligent Transportation Systems. Gran Canaria, Spain, 2015: 2788–2793. DOI: 10.1109/ITSC.2015.445
[30] ZOU Q, YAO X, ZHAO P, et al. Detecting Home Location and Trip Purposes for Cardholders by Mining Smart Card Transaction Data in Beijing Subway [J]. Transportation, 2016: 1–26. DOI: 10.1007/s11116-016-9756-9
[31] MA X, LIU C, WEN H, et al. Understanding Commuting Patterns Using Transit Smart Card Data [J]. Journal of Transport Geography, 2017, 58: 135–145. DOI: 10.1016/j.jtrangeo.2016.12.001
[32] HUANG L, LI Q, YUE Y. Activity Identification from GPS Trajectories Using Spatial Temporal POIs' Attractiveness [C]//2nd ACM SIGSPATIAl International Workshop on Location Based Social Networks. San Jose, USA, 2010: 27–30. DOI: 10.1145/1867699.1867704
[33] FURLETTI B, CINTIA P, RENSO C, et al. Inferring Human Activities from GPS Tracks [C]//ACM SIGKDD International Workshop on Urban Computing. Chicago, USA, 2013. DOI: 10.1145/2505821.2505830
[34] LIU X. Inferring Trip Purposes and Uncovering Travel Patterns from Taxi Trajectory Data [J]. Cartography & Geographic Information Science, 2016, 43(2): 103–114. DOI: 10.1080/15230406.2015.1014424
[35] LIU X, HE J, YAO Y, et al. Classifying Urban Land Use by Integrating Remote Sensing and Social Media Data [J]. International Journal of Geographical Information Science, 2017, 31(8): 1675–1696. DOI: 10.1080 / 13658816.2017.1324976
[36] YUAN N J, ZHENG Y, XIE X, et al. Discovering Urban Functional Zones Using Latent Activity Trajectories [J]. IEEE Transactions on Knowledge & Data Engineering, 2015, 27(3): 712–725. DOI: 10.1109/TKDE.2014.2345405
[37] YAO Y, LIU X, LI X, et al. Mapping Fine-Scale Population Distributions at the Building Level by Integrating Multi-Source Geospatial Big Data [J]. International Journal of Geographical Information Science, 2017, 31(6): 1220–1244. DOI: 10.1080/13658816.2017.1290252
[38] CASCETTA E. Transportation Systems Engineering: Theory and Methods [M]. Dordrecht, Netherlands: Kluwer Academic Publishers, 2001. DOI: 10.1007/978-1-4757-6873-2
[39] ABUHAMOUD M A A, RAHMAT R A O K, ISMAIL A B. Modeling of Transport Mode in Libya: A Binary Logit Model for Government Transportation Encouragement [J]. Australian Journal of Basic & Applied Sciences, 2011, 5(5): 1291–1296
[40] PRAVEEN KUMAR M, MALLIKARJUNA C. Mode Choice Modelling for Intercity Transportation in India: A Case of Guwahati to Five Metro Cities [J]. International Journal of Earth Sciences and Engineering, 2011, 4(6): 364-374
[41] BHAT C R. A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice [J]. Transportation Research Part B Methodological, 1995, 29(6): 471–483. DOI: 10.1016/0191-2615(95)00015-6
[42] HESS S. Advanced Discrete Choice Models with Applications to Transport Demand [D]. London, UK: Imperial College London, 2005
[43] MINAL, SEKHAR C R. Mode Choice Analysis: The Data, the Models and Future Ahead [J]. International Journal for Traffic and Transport Engineering, 2014, 4(3): 269–285. DOI: 10.7708/ijtte.2014.4(3).03
[44] GHAREIB A H. Evaluation of Logit and Probit Models in Mode-Choice Situation [J]. Journal of Transportation Engineering-Asce, 1996, 122(4): 282–290. DOI: 10.1061/(ASCE)0733-947X(1996)122:4(282)
[45] DOW J K, ENDERSBY J W. Multinomial Probit and Multinomial Logit: a Comparison of Choice Models for Voting Research [J]. Electoral Studies, 2004, 23 (1): 107–122. DOI: 10.1016/S0261-3794(03)00040-4
[46] BHAT C R, SARDESAI R. The Impact of Stop-Making and Travel Time Reliability on Commute Mode Choice [J]. Transportation Research Part B-methodological, 2006, 40(9): 709–730. DOI: 10.1016/j.trb.2005.09.008
[47] ABDEL-ATY M, ABDELWAHAB H. Calibration of Nested Mode Choice Model for Florida [R], Final Research Report, University of central Florida, 2001.
[48] GAUNDRY M J, DAGENAIS M G. The Dogit Model [J]. Transportation Research Part B-methodological, 1979, 13(2): 105–111. DOI: 10.1016/0191-2615 (79)90028-6
[49] HENSHER D A, TU T T. A Comparison of the Predictive Potential of Artificial Neural Networks and Nested Logit Models for Commuter Mode Choice [J]. Transportation Research Part E Logistics & Transportation Review, 2000, 36 (3): 155–172. DOI: 10.1016/S1366-5545(99)00030-7
[50] CANTARELLA G E, DE LUCA S. Modeling Transportation Mode Choice Through Artificial Neural Networks [C]///Fourth International Symposium on Uncertainty Modeling and Analysis, 2003. College Park, USA, 2003: 84–90. DOI: 10.1109/ISUMA.2003.1236145
[51] OMRANI H. Predicting Travel Mode of Individuals by Machine Learning [J]. Transportation Research Procedia, 2015, 10: 840–849. DOI: 10.1016 / j.trpro.2015.09.037
[52] LEE D, DERRIBLE S, PEREIRA F C. Comparison of Four Types of Artificial Neural Network and a Multinomial Logit Model for Travel Mode Choice Modeling [J]. Transportation Research Record Journal of the Transportation Research Board, 2018. DOI:10.1177/0361198118796971

Biographies

ZHAO Juanjuan (jj.zhao@siat.ac.cn) received the Ph.D. degree from Chinese Academy of Sciences, China in 2017. She is an assistant professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. Her research interests include big data processing, data privacy, and urban computing.

XU Chengzhong received the Ph.D. degree from the University of Hong Kong, China in 1993. He is the Dean of the Faculty of Science and Technology, University of Macau, China and a Chair Professor of Computer Science of UM. He was a Chief Scientist of Shenzhen Institutes of Advanced Technology (SIAT) of Chinese Academy of Sciences and the Director of Institute of Advanced Computing and Digital Engineering of SIAT. He was also in the faculty of Wayne State University, USA for 18 years. Dr. Xu's research interest is mainly in the areas of parallel and distributed systems, cloud and edge computing, and data-driven intelligence. He has published over 300 peer-reviewed papers on these topics with over 10K citations. Dr. Xu served in the editorial boards of leading journals, including IEEE Transactions on Computers, IEEE Transactions on Cloud Computing, IEEE Transactions on Parallel and Distributed Systems, and Journal of Parallel and Distributed Computing. He is the Associate Editor-in-Chief of ZTE Communication. He is IEEE Fellow and the Chair of IEEE Technical Committee of Distributed Processing.

MENG Tianhui received the Ph.D. degree in computer science from Free University of Berlin, Germany in 2017. He is currently an assistant professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China. His research interests include mobile edge computing, big data processing, blockchain and Internet of Things.