

Journal of Advanced Transportation

Emerging Data for Smart Transportation Management

Lead Guest Editor: Kun Xie

Guest Editors: Hong Yang and Yuanyuan Zhang





Emerging Data for Smart Transportation Management

Journal of Advanced Transportation

Emerging Data for Smart Transportation Management

Lead Guest Editor: Kun Xie





Guest Editors: Hong Yang and Yuanyuan Zhang



Copyright © 2021 Hindawi Limited. All rights reserved.






















This is a special issue published in "Journal of Advanced Transportation." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Associate Editors

Juan C. Cano , Spain
Steven I. Chien , USA
Antonio Comi , Italy
Zhi-Chun Li, China
Jinjun Tang , China

Academic Editors

Kun An, China
Shriniwas Arkatkar, India
José M. Armingol , Spain
Socrates Basbas , Greece
Francesco Bella , Italy
Abdelaziz Bensrhair, France
Hui Bi, China
María Calderon, Spain
Tiziana Campisi , Italy
Giulio E. Cantarella , Italy
Maria Castro , Spain
Mei Chen , USA
Maria Vittoria Corazza , Italy
Andrea D'Ariano, Italy
Stefano De Luca , Italy
Rocío De Oña , Spain
Luigi Dell'Olio , Spain
Cédric Demonceaux , France
Sunder Lall Dhingra, India
Roberta Di Pace , Italy
Dilum Dissanayake , United Kingdom
Jing Dong , USA
Yuchuan Du , China
Juan-Antonio Escareno, France
Domokos Esztergár-Kiss , Hungary
Saber Fallah , United Kingdom
Gianfranco Fancello , Italy
Zhixiang Fang , China
Francesco Galante , Italy
Yuan Gao , China
Laura Garach, Spain
Indrajit Ghosh , India
Rosa G. González-Ramírez, Chile
Ren-Yong Guo , China



Yanyong Guo , China
Jérôme Ha#rri, France
Hocine Imine, France
Umar Iqbal , Canada
Rui Jiang , China
Peter J. Jin, USA
Sheng Jin , China
Victor L. Knoop , The Netherlands
Eduardo Lalla , The Netherlands
Michela Le Pira , Italy
Jaeyoung Lee , USA
Seungjae Lee, Republic of Korea
Ruimin Li , China
Zhenning Li , China
Christian Liebchen , Germany
Tao Liu, China
Chung-Cheng Lu , Taiwan
Filomena Mauriello , Italy
Luis Miranda-Moreno, Canada
Rakesh Mishra, United Kingdom
Tomio Miwa , Japan
Andrea Monteriù , Italy
Sara Moridpour , Australia
Giuseppe Musolino , Italy
Jose E. Naranjo , Spain
Mehdi Nourinejad , Canada
Eneko Osaba , Spain
Dongjoo Park , Republic of Korea
Luca Pugi , Italy
Alessandro Severino , Italy
Nirajan Shiwakoti , Australia
Michele D. Simoni, Sweden
Ziqi Song , USA
Amanda Stathopoulos , USA
Daxin Tian , China
Alejandro Tirachini, Chile
Long Truong , Australia
Avinash Unnikrishnan , USA
Pascal Vasseur , France
Antonino Vitetta , Italy
S. Travis Waller, Australia
Bohui Wang, China
Jianbin Xin , China



Hongtai Yang , China
Vincent F. Yu , Taiwan
Mustafa Zeybek, Turkey
Jing Zhao, China
Ming Zhong , China
Yajie Zou , China

Contents



Demonstration of Participation Networks in Urban Transport Policy of Public and Private Sectors through Social Media: The Case of Bike-Sharing Pricing Strategy in China

Qian Ye , Xiaohong Chen, Hua Zhang , Junjie Cai, and Kaan Ozbay
Research Article (15 pages), Article ID 8881106, Volume 2021 (2021)





Predicting Critical Bicycle-Vehicle Conflicts at Signalized Intersections

Alireza Darzian Rostami , Anagha Katthe , Aryan Sohrabi, and Arash Jahangiri 
Research Article (16 pages), Article ID 8816616, Volume 2020 (2020)


Analysis of Electric Vehicle Charging Behavior Patterns with Function Principal Component Analysis Approach

Chenxi Chen, Yang Song, Xianbiao Hu , and Ivan G. Guardiola 
Research Article (12 pages), Article ID 8850654, Volume 2020 (2020)

Development of a Model for Evaluating the Coverage Area of Transit Center Using Smart Card Data

Hyunjun Hwang , Shin-Hyung Cho , Dong-Kyu Kim , and Seung-Young Kho 
Research Article (13 pages), Article ID 8819791, Volume 2020 (2020)


Person Detection for an Orthogonally Placed Monocular Camera

Pavel Skrabanek , Petr Dolezel, Zdenek Nemecek, and Dominik Stursa
Research Article (13 pages), Article ID 8843113, Volume 2020 (2020)

Heterogenous Trip Distance-Based Route Choice Behavior Analysis Using Real-World Large-Scale Taxi Trajectory Data

Yajuan Deng, Meiyi Li, Qing Tang, Renjie He, and Xianbiao Hu 
Research Article (16 pages), Article ID 8836511, Volume 2020 (2020)

The Importance of Cognitive and Mental Factors on Prediction of Job Performance in Chinese High-Speed Railway Dispatchers

Zizheng Guo, Jiaming Zou, Chuanning He, Xi Tan, Chongshuang Chen, and Guo Feng 
Research Article (13 pages), Article ID 7153972, Volume 2020 (2020)

Research Article

Demonstration of Participation Networks in Urban Transport Policy of Public and Private Sectors through Social Media: The Case of Bike-Sharing Pricing Strategy in China

Qian Ye ¹, Xiaohong Chen,¹ Hua Zhang ², Junjie Cai,^{3,4} and Kaan Ozbay³

¹Key Laboratory of Road and Traffic Engineering of the Ministry of Education, School of Transportation Engineering, Tongji University, Shanghai 201804, China

²National Maglev Transportation Engineering R&D Center, Tongji University, Shanghai 201804, China

³C2SMART Center, Department of Civil and Urban Engineering, Tandon School of Engineering, New York University, New York, NY 11201, USA

⁴Baidu Inc., Beijing 100085, China

Correspondence should be addressed to Hua Zhang; xiaohai_hua@tongji.edu.cn

Received 8 May 2020; Revised 9 June 2020; Accepted 15 April 2021; Published 27 April 2021

Academic Editor: Jaeyoung Lee

Copyright © 2021 Qian Ye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Social media has become a valuable platform that enables public and private stakeholders to participate and interact in various transport policies. Using a network-based perspective and a case study of bike-sharing pricing strategies in China, this paper aims to quantitatively characterize the pattern and structure of multi-stakeholders engagement networks. Furthermore, this paper also empirically examines the confirmation bias that might exist among participants. Dataset on retweets from the Chinese Twitter-Sina Weibo is collected. Results reveal two types of important actors with unequal roles in terms of information diffusion: the “network root” and the “network bridge.” The former is mainly comprised of organizations and influential individuals who dominate message sharing, whereas the latter is comprised of the general public with various occupational backgrounds who control the efficiency and the scope of information spreading. The result also reveals a hierarchical structure in both networks and a community gathering like-minded individuals. The empirical result also demonstrates the existence of echo chambers in the transport participation network of governments and enterprises. Most echo chambers operate such that organizations or influential individuals amplify the views of the general public with more critical viewpoints. These findings of this study can assist transport stakeholders in crafting more sustainable strategies based on the understanding of uneven patterns in online public participation. Furthermore, this study sheds insights on how social media could be used to facilitate the collection of diverse people’s opinions and the evaluation of multi-stakeholder engagement for major transport issues.

1. Introduction

Making appropriate policies or strategies is essential to the sustainable development of urban transport. Traditionally, transport policy-making processes are characterized by a top-down approach that the public seldom has access to, for example, hearings, expert consultants, or publications. It causes ineffective communication between citizens, private sectors, and administrative agencies as well as limited transparent governance [1]. In recent years, with the emerging of new mobility patterns and services, transport

governance may be undertaken by multiple stakeholders who have various interests, usually requiring interdisciplinary and multiagent cooperation [2]. Consequently, it highlights the differences in urban transport between the concept of government-driven sharing and that of collaboration among multi-stakeholders, i.e., market regulators, service providers, and transport system users.

Social media has been widely recognized as being able to facilitate the engagement of experts, private sectors, civic organizations, and the public as well as being able to create large and diffused networks of relationships [3, 4]. Due to the

potential role of social media in extracting peoples' real thoughts and addressing the complexity of stakeholder relations in the transport field, scholars have explored its applications on both bottom-up and top-down engagement in urban transport [5]. However, this research remains in the infant stage.

In recent years, more and more private stakeholders (i.e., platform operators, service providers) have offered personal and urban level mobility services and thus a method for opinion collection that can be significant to business promotion strategies. As such, the structure and characteristics of public participation might be different in public and private decision-making processes in the transport field. However, few scholars have investigated the policy-making participation network from a comparative perspective.

Most recently, a phenomenon of confirmation bias has arisen from concerns about decision-making or information processing: that is, the tendency to search for, interpret, and recall information in a way that is consistent with existing beliefs [6]. In social media, it is defined as an echo chamber or filter bubbles [7]. The presence of echo chambers leads to information barriers around people and, thus, causes difficulties in sharing and understanding diverse viewpoints and even reaching a solution. As such, it is necessary to confirm the existence and the degree of confirmation bias on the collected social media data before forming a conclusion on transport policy issues. However, little evidence has concerned echo chambers in transport policies and their influence on the public and private sectors.

For those reasons, we considered it important to present and analyze the pattern of participation networks in transport policy-making led by governments (GCPN) and enterprises (ECPN). This study also aims to empirically examine the existence of confirmation bias among participants in terms of personal attitude diffusion. Two topics on bike-sharing pricing strategies resulting from government regulations and business strategies are empirically studied, by collecting the retweet data from the Chinese Twitter-Sina Weibo. The specific objectives are (i) to illustrate the pattern and characteristics of the citizen involvement in transport policies by public and private sectors, (ii) to simplify the structure of participation networks and identify influential actors/groups with key roles in information dissemination, and (iii) to empirically examine the degree of confirmation bias in transport policy participation networks. The results can help multiple stakeholders to improve transport policies from a perspective of the nature of online public engagement. Furthermore, results provide insights into the design of evidence-driven policy research to avoid the potential bias when using social media as a platform to gather opinions and needs of people from different backgrounds concerning transport issues.

2. Literature Review

2.1. Online Public Participation in the Transport Policy-Making. Online platforms (e.g., social media, blogs, and websites) have become a vital channel for public involvement that may be supplementary to traditional or face-to-

face engagement opportunities [8]. Increasing theoretical efforts and implementations have been devoted to the potential of social media in understanding the complex interactions between transport stakeholders and the public in the transport domain [9, 10]. For example, transport policymakers and providers in the USA, like MTA Transit and Metro-North Railroad in New York City, and DOTs in Alabama, California, West Virginia, etc., use Facebook and Twitter to inform travel information or specific events as well as to connect the feedback from the residents, so that to facilitate project improvements and misinformation corrections [8, 11]. Furthermore, with the support of web-based tools/toolkits and text mining techniques [12], ridesharing agencies can optimize ridership, timeliness, efficiency, and safety and also reveal the level of satisfaction with ride-sharing services by monitoring social media [13, 14]. Another example is the adjustment of the No. 55 bus route in Shanghai. The transport agencies and local magazines disseminated the message about the abolition of the bus route on Weibo. Since a large number of citizens commented on this message and expressed their nostalgia for the bus route, the local government was pushed to interact with citizens and completely change the initial plan [15].

To date, many methodologies like network science are developed to demonstrate the complex participation networks involving multiple stakeholders [16]. Social Network Analysis (SNA) is a method proposed by sociologists to measure the positions and structural characteristics of relationships between social actors [17, 18]. The main research topics of SNA in the transport field are summarized into three categories: megaproject collaboration and management [19, 20], multimodal transport network analysis [21, 22], and participatory decision-making. Of very few studies using SNA on transport policy issues, Dörry and Decoville study the situated multilevel governance structure of cross-border public transport provision by taking the respective cross-border policy network in the Luxembourg metropolitan region as a case [23]. Sun et al. use the SNA method to refine multi-actor multicriteria analytical results and further evaluate various low-carbon transport policies [24].

According to the study conducted by Weiss et al. and Mills et al., SNA can highlight which actors are important for influencing policies or actions, or for facilitating information and knowledge transmission [25, 26]. Therefore, this study applies the SNA method to simplify the structure of participation networks in terms of public and private policy-making.

2.2. Confirmation Bias on Social Media and Real-World Policy Networks. Confirmation bias is a psychological phenomenon where humans tend to seek information adhering to their preexisting opinion [27], which has become a widely debated issue at both the scientific and political levels [28]. Recent studies have shown that information bias is dominant in the social media environment [29, 30], where individuals are surrounded by people whose viewpoint agrees with their own, also known as echo chambers or filter

bubbles [7]. Major social media like Facebook and Twitter presents clear-cut echo chambers on controversial topics [31]. This configuration could cause the spreading of bias knowledge [32, 33], hamper the democratic deliberative process by altering the way facts are perceived [34], and even lead to a loss of common understanding [35]. Several studies develop the theoretical definitions and mechanisms of echo chambers [31]. For instance, Jasny, Waggle, and Fisher propose a more formal operationalization of echo chamber components, which comprises two processes of an “echo” and “chambers” and at least three types of actors—speakers, receivers, and mediating actors [36]. Cinelli et al. introduce a definition and method of echo chambers based on the co-existence of two main ingredients: opinion polarization concerning controversial topics and homophilic interactions between users [31].

From the policy network perspective, echo chambers are questioned to affect evidence-based policy-making as they can hinder policy learning and consensus building [28]. It has been demonstrated that organizations and individuals tend to obtain information from those with similar beliefs to theirs while ignoring information from sources that challenge their beliefs [36, 37]. Recent studies also provide empirical evidence of the effects of echo chambers on policy networks using survey data. For example, Jasny et al. compare two datasets on the US climate politics at two points in time and find that echo chambers continue to play an important role in the information sharing among policy elites working on climate issues [38]. Wagner and Ylä-Anttila find that echo chambers are present in the relationship between the information-seeking behaviors of policy actors in Irish policy networks; also, they reveal that policy forums could help alleviate the effects of echo chambers if organizations with different views were to participate and to use the opportunities that forums provide to learn from those outside their networks [28].

Due to the complexity of disciplinary issues emerging in shared transport and the essentiality of information sharing among multiple stakeholders, the phenomenon of echo chambers is potentially influencing the policy-making in urban transport. Therefore, there is a need to test the existence of echo chambers in public and private policy-making and further to emphasize its applicability and effectiveness in the analysis of various transport policies.

3. Background and Research Methods

3.1. Case Information on Bike-Sharing Pricing Policies in China. The dockless bike-sharing service was introduced to China since the establishment of the Ofo platform in 2014. The system consists of government regulators, platform operators, bike associations, and bike-sharing users [39]. The number of bike-sharing users increased significantly within the first two years. However, due to the lack of sustainable business strategies and management, leading bike-sharing enterprises like Ofo have faced serious financial problems since the end of 2018. Millions of riders who placed deposits demanded their bicycle refunds, causing widespread concerns and debates among the public and government

authorities on deposit-related issues of emerging travel services in China [40].

In response, the Ministry of Transport (MOT) in China formally has included the deposit issue in government supervision and published tighter controls soon. In March 2019, an administrative measure of user deposit in emerging transport was issued to gather advice from the public, followed by the final version of the implementation guidance that was released in May 2019 [41]. This document specifies the behaviors of service providers including the schedule of refunding user deposits and the way of deposit management. Soon after, several bike-sharing enterprises including Bluegogo, Mobike, and Hello announced their pricing increase strategies. More and more enterprises also employed a higher level of bike-sharing fares in the next months. The outline of bike-sharing pricing strategies in China is shown in Figure 1.

Therefore, this study chooses these two topics on Weibo to understand the citizen participation in public and private decision-making, corresponding to the fact that “Chinese governments have issued an administrative measure of user deposits in emerging transport services” and “many enterprises have launched the price increase in bike-sharing services.”

3.2. Research Methods. The analytical framework of this study includes the collection of retweet data on Weibo, SNA-based measurement, and echo chamber examination, which is depicted in Figure 2.

3.2.1. Data Collection. A data set of Weibo retweets on the abovementioned two topics was collected on July 8, 2019. It consists of two sections: messages/retweets and information of participated users. A snowball sampling method was used to complete the sample of participants and their messages in this study. This data collection method has been employed in the analysis of social networks and social dynamics [42]. For each topic, the initial samples were top-50 most-read messages on Weibo (called the initial messages). Then, messages forwarded from the initial messages by users were picked (called the first layer of retweets). Similarly, retweets of deeper layers were further searched until no new retweets were founded. The data set of participant information was gathered from their homepage on Weibo addressed by their usernames shown in the retweet records. Available features such as job, education, and identity verification were extracted. The verification setting on Weibo enables businesses, celebrities, etc. to show their credibility and trust, generally correlated with their numbers of followers and influence. The participants were correspondingly divided into three categories: influential individuals, organizations/companies, and the general public. Finally, a total of 2335 and 2845 participants who published 2469 and 2938 Weibo messages related to two topics were included in this study.

The next step was to reconstruct the forwarding/forwarded relationship among participants by parsing the content of retweets. Through identifying unique user IDs

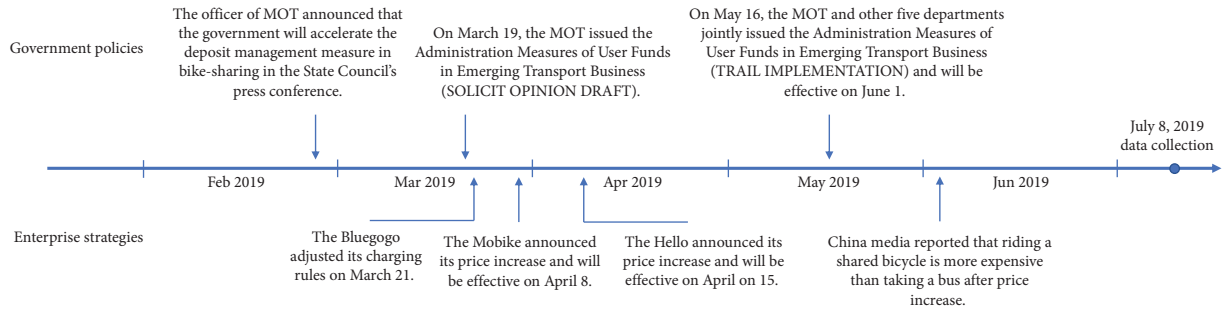


FIGURE 1: Overview of deposit-related policies of bike-sharing in China.

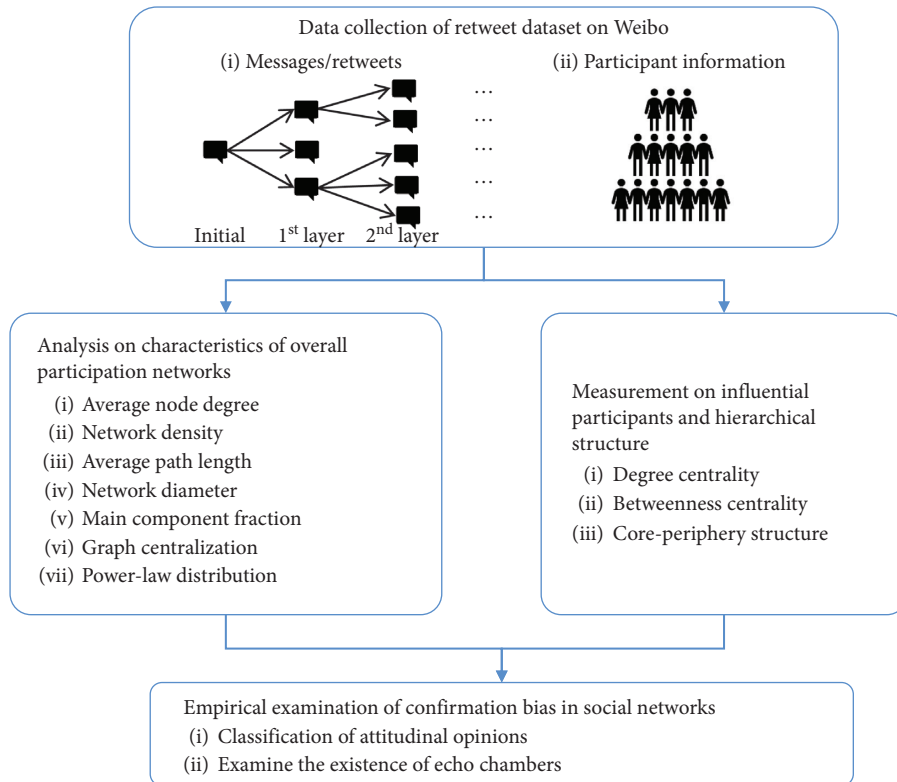


FIGURE 2: Research analysis framework in this study.

and the retweet identifier (“//@” on Weibo) in the message dataset, direct/indirect interactions and corresponding information flow started from the initial participants were obtained. To map and analyze the data sets, an adjacency matrix, a square matrix used to represent a finite graph, was used to represent the relational data among the pairs of vertices (participants) [43]. The number of each cell $r_{i,j}$ in the matrices was denoted as 1 for two participants (i.e., i in the row and j the column) who were directly linked in the forwarding/forwarded structure and denoted as 0 otherwise. A rule was applied here: that is, simplifying the multiple forwarding/forwarded between the same pair of users into one. Thus, the value of $r_{i,j}$ was binary, in which edges among participants were dichotomized as 1 or 0.

3.2.2. Network Classification and Measurement. Given this unweighted directed graph, the network was formalized as a collection of participants as nodes and their forwarding/forwarded relationships as edges using the SNA method. For example, there was a directed edge from actor A to actor B when B forwarded a message from A, which is consistent with the direction of information flows within networks. As this study focuses on comparing the participation networks of public and private sectors, datasets were further classified into two examined networks: citizen participation network led by governments (represent as GCPN) and enterprises (represent as ECPN). As such, a total of 2390 edges (nonzero cells $r_{i,j}$) in GCPN and 2873 edges in ECPN were analyzed. Descriptive analysis characterizes both the overall structure

and the individual performance of network nodes. The indicators used are as follows [44]:

- (i) Average node degree, which refers to the average number of edges per node in networks. It includes in-degree and out-degree which indicate the number of forwarded messages to and from each participant.
- (ii) Network density, which measures the tightness of nodes that are calculated as the proportion of edges that are existing in the network to the largest number of possible edges. Its values vary between 0 and 1. The greater the density of networks, the more the contacts among participants.
- (iii) Average path length, which refers to the average number of steps along the shortest paths for all possible pairs of nodes. It measures information efficiency on networks.
- (iv) Network diameter, which refers to the shortest distance between the two most distant nodes in networks. It determines an upper bound of the distance between any pairs of participants who can exchange information.
- (v) Main component fraction, which refers to the proportion of main components (or giant components). It is a node group with the largest number of interconnected nodes that emerged in networks.
- (vi) Graph centralization, which refers to the degree to which the centrality of the most central node exceeds the centrality of all other nodes in overall networks. It varies between 0 and 1, which quantifies the variation of individual nodes' importance and the distribution of node positions in networks.
- (vii) Power-law distribution, which refers to the cumulative probability of nodes whose degree is larger than or equal to k (represented as $D(k)$), to assess the unevenness of specific network indicators (e.g., out-degree). The distribution is formulated as follows [45], where α is the law's exponent and C is a constant:

$$\begin{aligned} D(k) &= C \cdot k^{-\alpha}, \\ \ln D(k) &= \ln C - \alpha \ln k. \end{aligned} \quad (1)$$

Concerning the different roles of individual nodes and the hierarchical structure in each participation network, indicators of degree centrality and betweenness centrality, as well as k -core decomposition, were selected to identify influential participants/groups in terms of information dissemination. Degree centrality (DC) refers to the number of edges that a node has, which is a primary measure of the importance of nodes [46]. In a forwarding/forwarded network, a relatively high out-degree centrality indicates that the corresponding actor has more channels of information exchanges with other participants and is thus at a more central position in the entire network (i.e., "network root"). Betweenness centrality (BC) is another measure to capture

essential nodes based on the shortest paths. It is calculated as the number of these shortest paths that pass through target nodes [47]. In terms of participation networks, a specific person with a high BC means that they played a significant role (e.g., "network bridge") in allowing information to travel from one part to another in the entire network.

The k -core decomposition is an alternative degree-based method to conduct hierarchical structure analysis, graph visualization, and community detection in complex networks [48]. The k -core indicates the maximal subgraph in which every node is adjacent to at least k nodes. An integer index, k_s , is assigned to each node to represent its location in networks; the largest value is denoted as k_{\max} (called k -core index). Each layer that contains nodes of degree k is called k -shell. In this sample, participants with low or high k_s are correspondingly located at the periphery or the core, also indicating the importance of nodes in the entire network. It should be noted that nodes with a high degree may or may not have a high k_s .

3.2.3. Echo Chamber Identification Algorithm. This study employs an operationalization of the components of echo chambers proposed by Jasny, Waggle, and Fisher [36]. Given the network formalism in Section 3.2.1 (represented in Figure 3(a)), two characteristics of echo chambers in participation networks are specified. An "echo" occurs when two participants share the same opinion on the policy. The attribute is indicated as the shading of the nodes in Figure 3(b). A "chamber" is formed when the information from the same source eventually reaches the same target via multiple different paths. For example, a transitive triad is the smallest example of a "chamber" where messages pass from participant A to participant C through a directed edge, but also indirectly through participant B as B forwards the message from A and then C forwards the message to B (see Figure 3(c)).

Thus, echo chambers can be demonstrated once the chamber is configured around participants holding the same opinions (see Figure 3(d)).

Considering the feature that retweets are more likely to be passed from a few network elites to a large number of users, this study developed a more effective identification algorithm to find echo chambers quickly from online participation networks.

- (i) Step 1: search participants that forward information from multiple sources, for example, find out C who receives messages from at least two sources in Figure 3
- (ii) Step 2: select the structure where connections between any two sources exist, that is, to find out a like-shaped transitive triad
- (iii) Step 3: test the uniqueness of participants in the structure, and then find out the transitive trail that has three distinct participants—a "chamber"
- (iv) Step 4: determine the presence of an "echo chamber" by checking whether the participants in chambers hold the same viewpoints

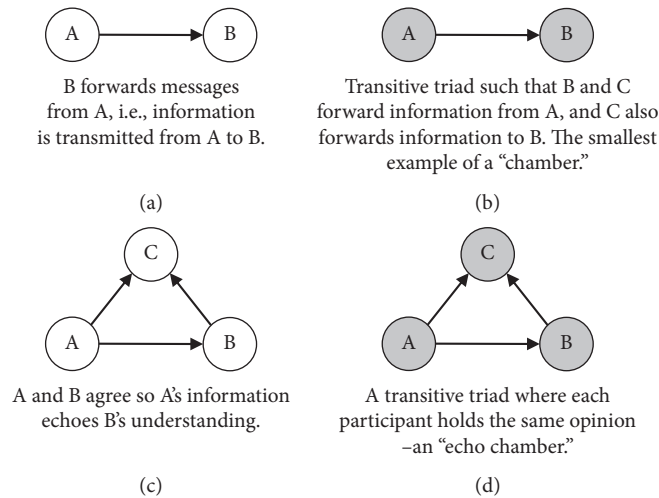


FIGURE 3: Structural and attitudinal components of an echo chamber in participation networks.

4. Communication Patterns in Transport Policy-Making

4.1. Descriptive Analysis of Network Characteristics of Participation Networks. Descriptive analysis of the network of GCPN and ECPN was firstly calculated with UCINET 6.665 and then visualized through its embedded drawing package *NetDraw* [49]. To present the key structure in two studied networks, Figure 4 visualizes the simplified networks where marginal nodes with low out-degrees (i.e., have few connections with other network actors) are removed. Different node sizes are proportional to out-degrees of the nodes (i.e., the number of forwarded messages from each participant), and different node shapes are used to distinguish the identity of participants in transport policy networks, with squares indicating organizations/companies, diamonds indicating influential individuals, and circles indicating the general public.

The network of GCPN engages 157 participants who include 49 organizations/companies and 13 influential individuals (see Figure 4(a)), while there is a total of 140 participants and a smaller proportion of verified participants in ECPN, i.e., 24 organizations/companies and 3 influential individuals (see Figure 4(b)). The results show that topics related to bike-sharing business strategies were more attractive to the general public than topics related to government regulations. In other words, in this case, companies can use the Weibo platform to obtain more opinions and feedback from ordinary users from the online communication and participation process. Besides, a similar pattern can be discovered in both networks of GCPN and ECPN in that the entire network was sparse whereas individuals were centralized within several groups (i.e., two major groups in GCPN and three major groups in ECPN). In general, these findings reveal the characteristic of more open and decentralized policy participation than that of traditional top-down policy-making [15].

The network structure statistics were then calculated to quantitatively characterize the pattern of networks of GCPN

and ECPN (see Table 1). The values of average degree and network density are very close in the two studied networks, which reach 1.02 and $4.43e-4$ in GCPN and 1.01 and $3.57e-4$ in ECPN, respectively. It reveals that both of the network participants were weakly connected to others and the majority shared information in a single path. The values of average path length in GCPN and ECPN are around 1. Additionally, the values of network diameter are 4 in GCPN and 5 in ECPN, and the main component fraction reaches approximately 100% in two networks. The result suggests that the entire participation network on the issue of bike-sharing pricing policies was a giant component, and the depth of information dissemination reached around five layers away from the initial messages. It demonstrates the formation of virtual communities that were not limited by geographical and time constraints, and these groups played a significant role in citizen participation and communication in terms of bike-sharing public and private policy-making. Furthermore, the value of the out-degree graph centralization is 0.35 in GCPN and 0.29 in ECPN. We would conclude that there was a tendency for the networks to develop around certain prominent actors, which also can be observed from the network visualization shown in Figure 4. In other words, the power of participants varied rather substantially, and the positional advantages were rather unequally distributed in these two examined networks.

4.2. Uneven Power and the Different Positions of Participants.

To investigate the tendency of a substantial amount of concentration in the network of GCPN and ECPN, two indicators were selected to measure the power of each participant (see Figure 5): the number of followers (denoted by D^f), which refers to the cyber power of actors on Weibo; and the number of messages forwarded (out-degree, denoted by D^{out}), which refers to the actors’ participation level.

The result shows that the cumulative distributions D^f (see Figures 5(a) and 5(c)) and the cumulative out-degree distributions D^{out} (see Figures 5(b) and 5(d)) decayed as a

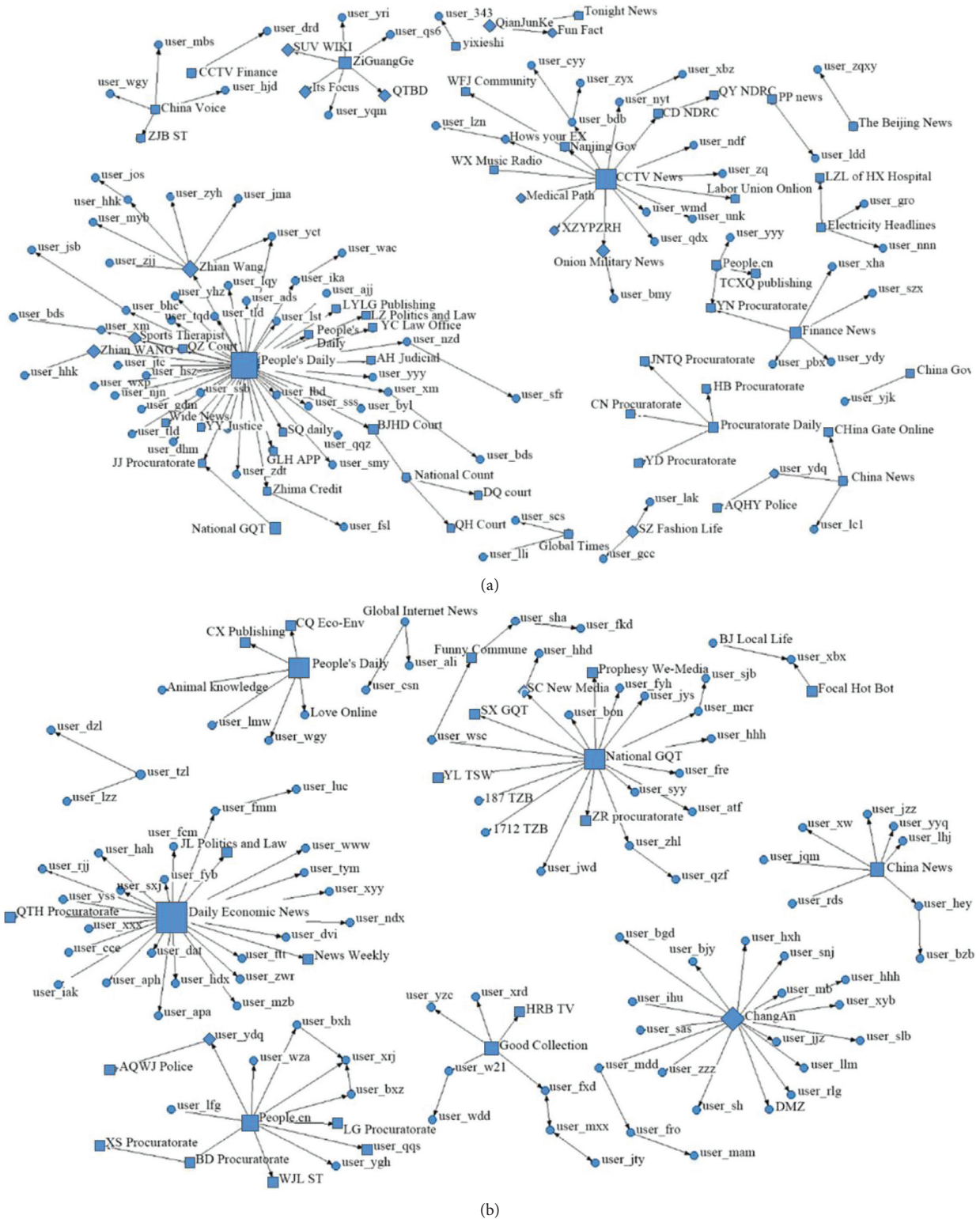


FIGURE 4: Simplified network mapping of (a) GCPN and (b) ECPN.

power-law scaling in GCPN and ECPN, suggesting the scale-free characteristic [50]. The pattern of D^f indicates that the cyber power of participants was uneven in terms of information diffusion on the topics of bike-sharing pricing policies. The distribution of D^{out} suggests that the influence

of various participants was significantly different because the messages of some actors were forwarded a greater number of times than those of other actors. That is, some participants probably played key roles and influenced others as the “networked power” (prominent participants) [51] in this

TABLE 1: Statistics of network characteristics.

Indicators	GCPN	ECPN
Average degree	1.02	1.01
Network density	$4.43e-4$	$3.57e-4$
Average path length	1.16	1.07
Network diameter	4.00	5.00
Main component fraction (%)	100.00	99.89
Out-degree centralization	0.35	0.29

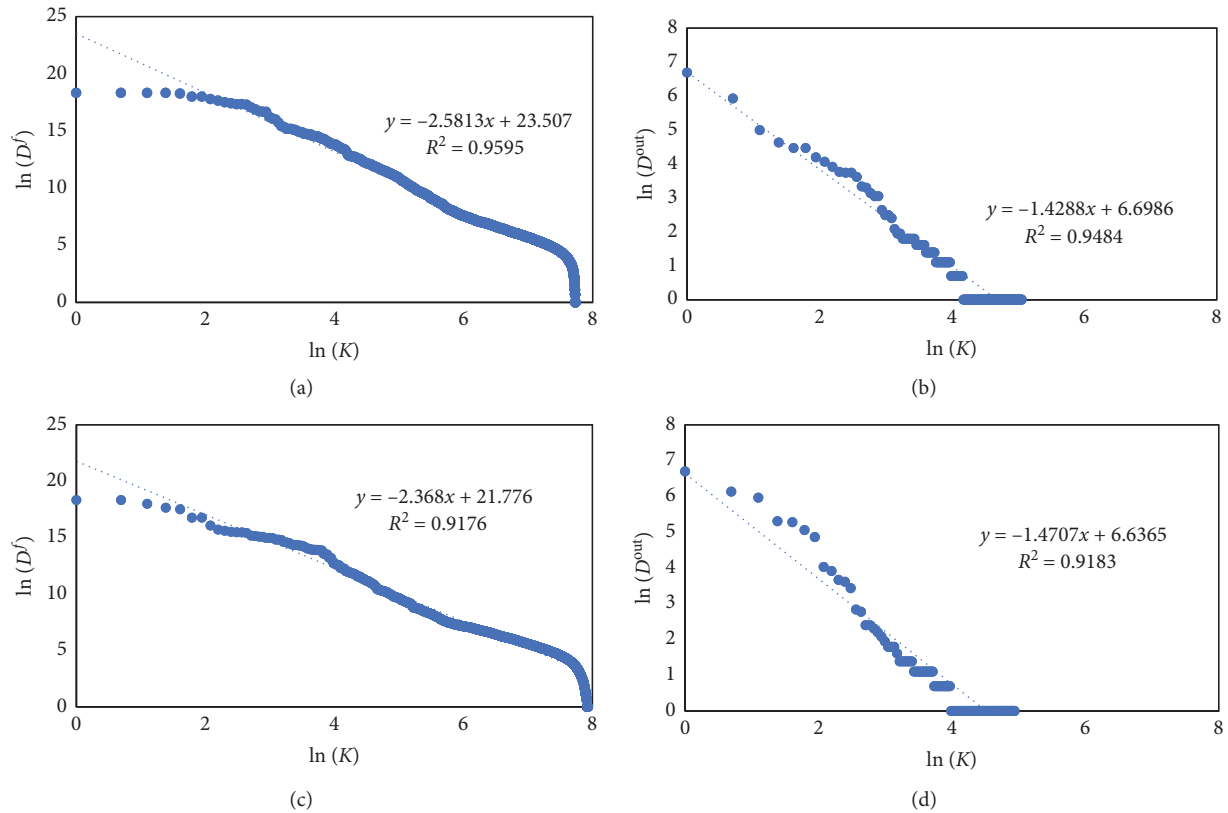


FIGURE 5: The distribution patterns in log-log graphs: (a) the number of followers (D^f) and (b) the number of messages forwarded (out-degree (D^{out})) in GCPN, (c) the number of followers (D^f), and (d) the number of messages forwarded (out-degree (D^{out})) in ECPN.

case. Taken together, these results provide quantitative evidence on the existence of the unequal power in information sharing of transport policies among various participants in both two studied networks.

5. Networked Power and Their Roles in Information Dissemination

5.1. Identifying Power Actors Using Centrality Measurements. To further identify the networked power in forwarding/forwarded networks of GCPN and ECPN, out-degree centrality and betweenness centrality were calculated. To make the results clear, the 15-top ranked participants in out-DC and BC lists were selected to analyze their identities and occupational backgrounds (see Figure 6).

In Figure 6(a), organizations/companies and influential individuals occupy the top-15 ranked positions of out-DC in

GCPN and ECPN, while a more proportion of the general public is presented at the top-15 in BC of both networks. This result suggests that verified actors who have the cyber power on Weibo participated actively and were located at the central position (i.e., the “network root”) in policy and private policy-making of bike-sharing. In contrast, the result of BC shows that the majority is the general public in this bike-sharing policy network; they played as the “network bridge.” The results are shown in Figure 6(b) to help to investigate the background of these influential participants. In terms of “network root,” both networks in GCPN and ECPN reveal a similar trend; that is, the majority were media companies and journalists; whereas the “network bridge” in this case were of diverse occupational backgrounds, such as officers from public agencies, academia excepts, and the media.

As the conclusion proposed by Berry [52], it is difficult for views based on a social network to play an important role in

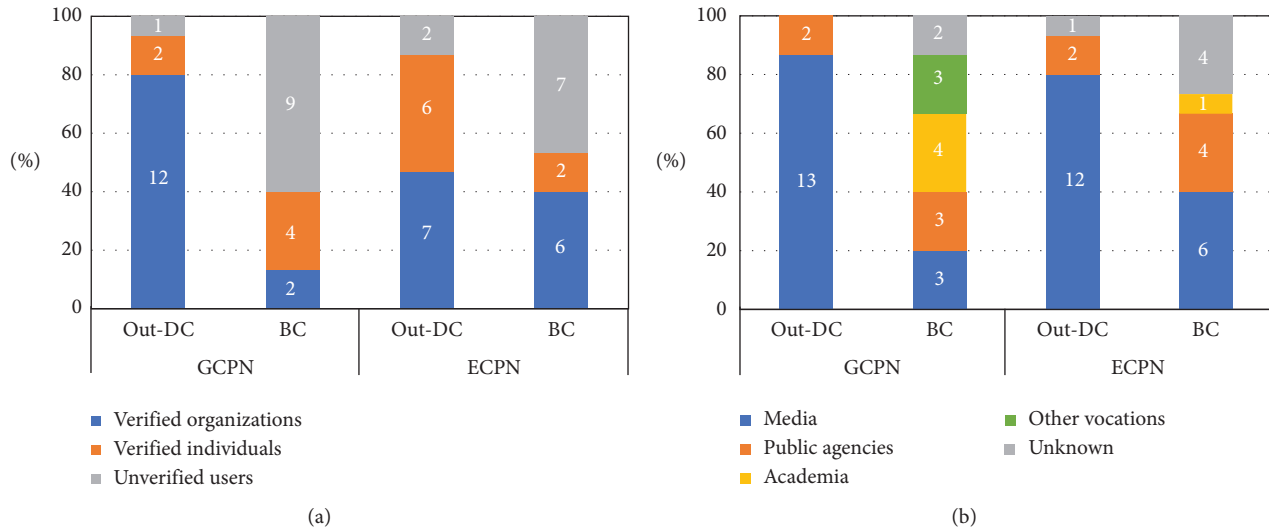


FIGURE 6: Fifty-top ranked participants aggregated by (a) verification identity and (b) occupational background.

decision-making processes without effective organization and cooperation. In the bike-sharing policy networks, it probably could be explained that organizations/companies and influential individuals established a large number of connections with others in the online discussion, while more diverse social actors were involved to help the information travel further within the virtual community. Therefore, there is a need for transport stakeholders to investigate and understand the needs of people who play different roles and have diverse backgrounds in public and private policy-making.

5.2. Exploring the Hierarchical Structure via K -Core Decomposition. The results of k -core decomposition reveal the hierarchical structure and the subgroups containing influential participants, as visualized in Figure 7, where nodes located at the most peripheral layer (0-shell) are removed. The colors of nodes are used to distinguish the different layers of k -shell, red indicating nodes located in 3-shell and green indicating nodes located in 2-shell. The k -core index of GCPN and ECPN is 3 and 2, respectively. There are 15 nodes in 3-shell and 81 nodes in 2-shell of GCPN, while there are 61 nodes in 2-shell of ECPN, which indicates that a small number of participants (i.e., 4.11% in GCPN and 2.14% in ECPN) were tightly connected in the networks.

As seen in Figure 7, the network of GCPN presents higher robustness than that of ECPN. The partition of GCPN had a three-layer structure (see Figure 7(a)), its 2-shell was denser with nodes tightly connected in multiple paths, whereas the 2-shell in ECPN is relatively separated or isolated (see Figure 7(b)). The result suggests that the participation network led by governments provided more flexible channels for actors to receive or transfer messages than that led by bike-sharing enterprises. Correspondingly, it could be explained that the aggregated opinions collected from the network of GCPN were closer to the thoughts of the majority on Weibo, while opinions generated from ECPN were probably influenced by their virtual communities.

Concerning the network of GCPN (see Figure 7(b)), there was a community that contains fifty important nodes in the core position. As the statement proposed by Zhao et al. [15], an important characteristic of the social society is that a group or community with similar values can be created and represented a collective identity. To validate the finding in this case, the content of messages and backgrounds related to these fifty participants were further investigated. There were 3 organizations/companies and influential individuals located at the center; the other 12 were the general public; they forwarded messages from three actors without posting any comments. The results reflect the same phenomenon found in Zhao et al.'s study; namely, the participants located in this community probably held the same or similar opinions on public policy-making related to bike-sharing. Another finding is that most of the participants within this 3-shell structure do not appear at the top-15 list of out-degree centrality and betweenness centrality, which indicates that participants may be of a high k_s but a low out-DC or BC. It could be explained that the core-periphery structure is another measurement of the importance of nodes focusing on the community level. Thus, it is suggested to employ various indicators or measurements to evaluate the structure and characteristics of policy networks.

6. Empirical Evidence of Confirmation Bias in Transport Policy Networks

6.1. Analysis of the Personal Viewpoints on Weibo. All the messages in two networks of GCPN and ECPN were reviewed and classified into three emotional opinions based on positive, negative, and neutral attitudes. Most participants held one of three viewpoints (see Figure 8):

- (i) The majority (over the half) forwarded the messages showing a neutral attitude or without posting any comments, i.e., 62.4% in GCPN and 55.0% in ECPN.

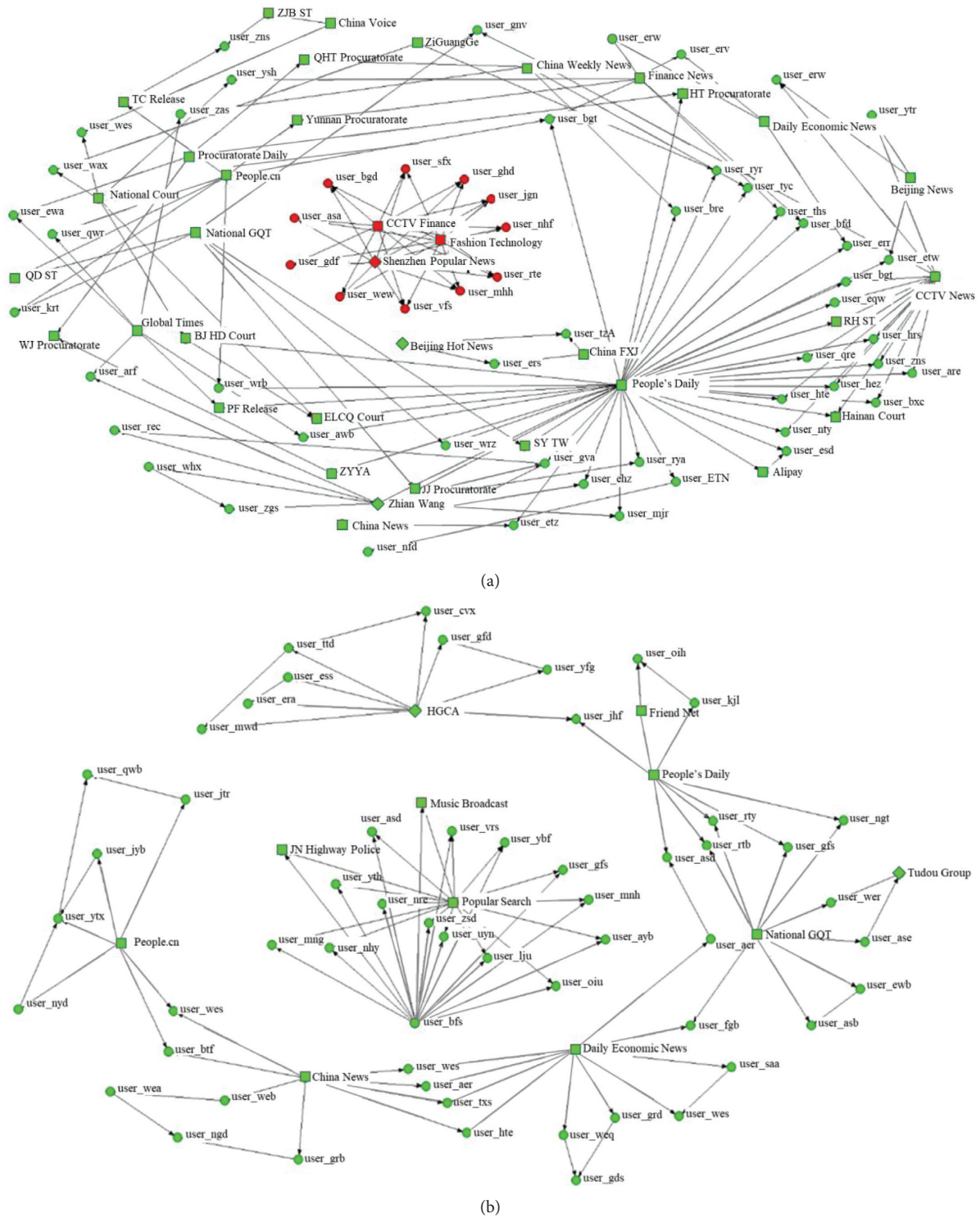


FIGURE 7: The core-periphery structure measured by K-core decomposition in (a) GCPN and (b) ECPN.

They played a significant role in transmitting their information in the communication process.

- (ii) Approximately 20.3% of the participants in GCPN expressed their pessimistic outlook on the bike-sharing pricing regulations implemented by governments. They worried that whether the

policy would be beneficial for refunding due to the dilemma that enterprises faced serious financial problems. At the same time, 33.7% of the participants in ECPN were against enterprises' increasing prices; they held that it is unreasonable to ride bikes with paying a higher fare that is

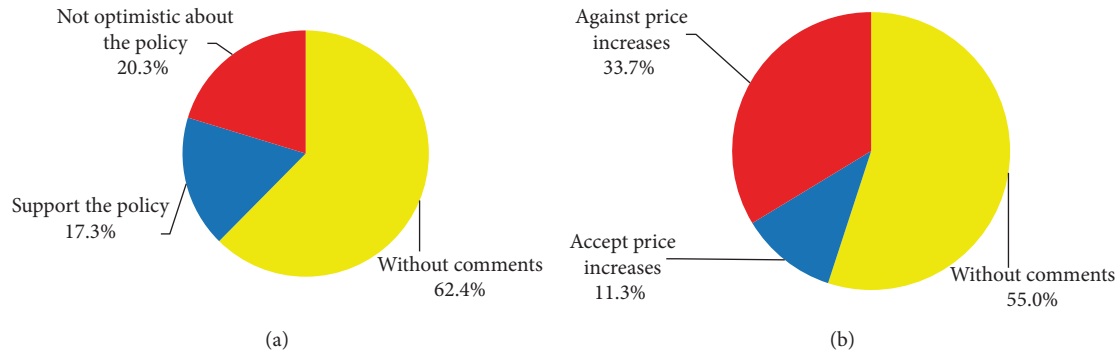


FIGURE 8: Different opinions in (a) GCPN and (b) ECPN.

even more expensive than that of subway or buses.

- (iii) A few of them showed their support or approval of the issue. For example, 17.3% of participants in GCPN appreciated the efforts paid by governments to regulate the bike-sharing industry and thus protect customers' rights. In ECPN, only 11.3% of the participants accepted the price increase of bike-sharing enterprises.

In general, the distribution of opinions reveals an important characteristic of online public engagement: that is, the majority will not express emotional views on public and private decision-making for major transport issues. The percentage of people holding negative opinions is roughly larger than that of positive opinions in the two examined networks.

6.2. Examination of the Presence of Echo Chambers in Transport Policy Networks. Results corresponding to each step proposed in the identification algorithm are listed in Table 2. It shows that there are five echo chambers and one echo chamber found in GCPN and ECPN, respectively. Generally, the result provides empirical evidence of the existence of echo chambers in participation networks of transport decision-making of public and private sectors.

However, the result shows that the proportion of echo chambers within entire participation networks was quite low, roughly at the same level in GCPN and ECPN (i.e., 3% and 2%). It suggests a relatively low degree of the impact of confirmation bias on online communication toward this bike-sharing case. Another finding is that over half of the chambers (i.e., 63%) in GCPN were satisfied with the attitudinal components while very few (i.e., 14%) in ECPN. It could be explained that due to the professional nature of transport governance, the general public is usually limited in participating in the discussion and expression of government policies. Thus, the participants in echo chambers of ECPN were more likely to agree with the views of influential actors, probably by the means of forwarding their messages without comments on social media.

In contrast, the transport service fare is the topic more closely to people's daily life; the business strategy adopted by

bike-sharing enterprises attracts more debates among residents. Correspondingly, the echo chamber provided a space needed for viewpoint exchange and free debates on the Internet.

The locations of echo chambers discovered in networks of GCPN and ECPN are visualized in Figure 9. The legend of node colors is changed here to indicate the opinions of each participant, which consists of the legend in Figure 8. All of the echo chambers are distributed in 2-shells; in other words, these participants involved in echo chambers were important actors in the hierarchical structure of transport policy networks as well.

As shown in the figure, there were three echo chambers configured around neutral attitudes (labeled as (i), (ii), and (v)), two echo chambers with negative attitudes (labeled as (iii) and (iv)) in GCPN (see in Figure 9(a)), and one echo chamber configured around neutral attitudes in ECPN (shown in Figure 9(b), labeled as (i)). Additionally, the mechanism of involved participants' identities and the path that echo chambers operate can be observed in the figure. In Figure 9(a), the echo chamber of neutral attitudes in GCPN existed in the information flow between organizations/companies (e.g., court departments) or from organizations/companies (e.g., nationwide news media) to the general public, whereas the echo chamber of negative opinions was presented in the message diffusion from influential individuals (e.g., famous journalists) to the general public. In contrast, as seen in Figure 9(b), an echo chamber with neutral attitudes was configured around an organization/company (e.g., nationwide news media) and two general participants. Furthermore, there is a phenomenon that happened in GCPN; namely, echo chambers in policy networks of this case were overlapped among participants in two echo chambers; see examples with labeling (i) and (ii) in Figure 9(a).

The above results could be explained; the networked power with greater influence and readership on social media was probably becoming the actors that echo/amplify others' viewpoints and understanding of major transport issues. Furthermore, according to this case, echo chambers were more likely to be configured around participants with pessimistic or neutral attitudes on bike-sharing pricing policies.

TABLE 2: The summary of echo chamber identification in participation networks.

Count	GCPN	ECPN
Step 1: multifoward participants	75.00	57.00
Step 2: like-shaped transitive triads	10.00	20.00
Step 3: chamber structure	8.00	7.00
Step 4: echo chambers	5.00	1.00
Percentage (%)	GCPN	ECPN
Echo chambers within networks	0.03	0.02
Echo chambers within chambers	0.63	0.14

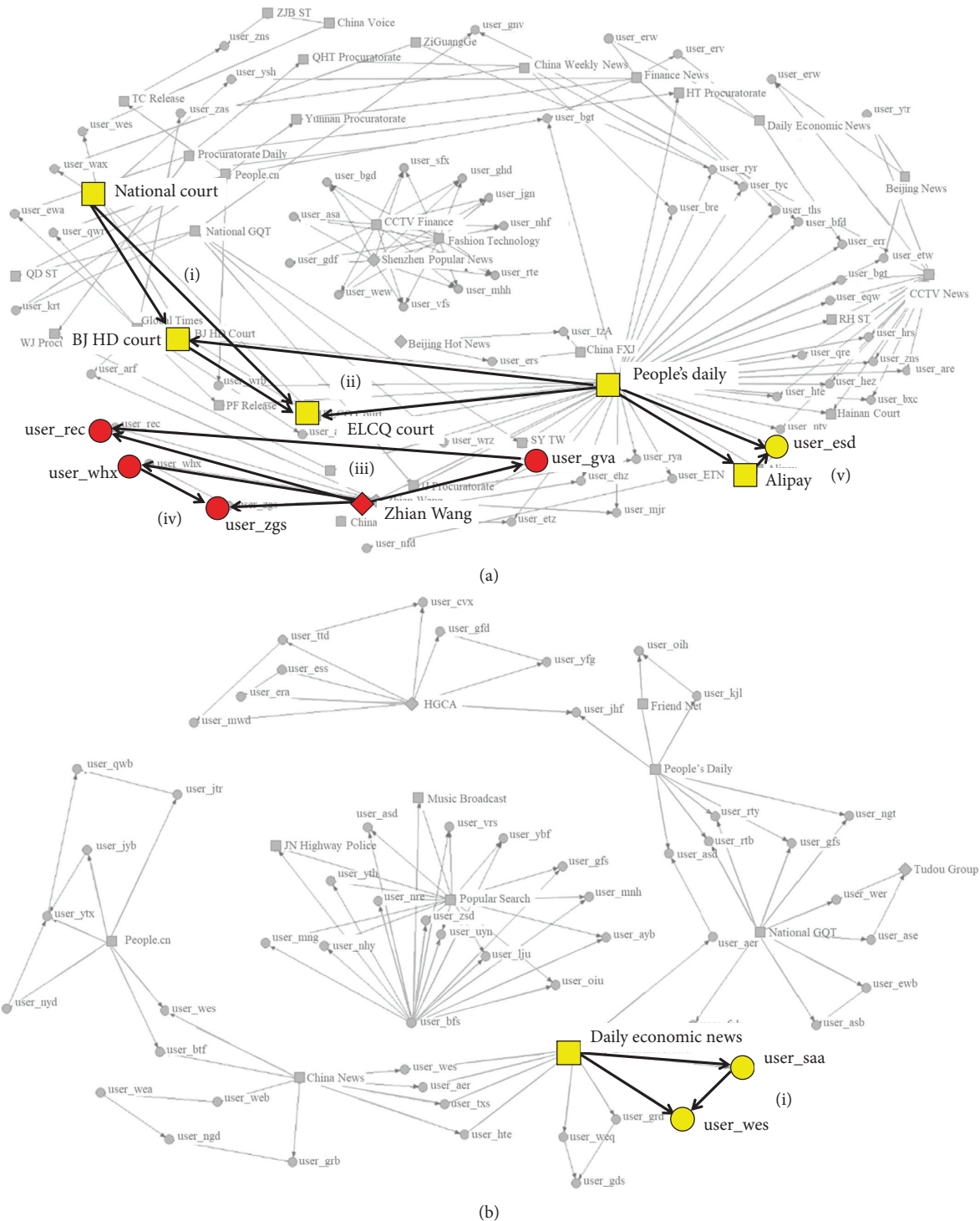


FIGURE 9: Illustration of the existence of echo chambers in the structure of (a) GCPN and (b) ECPN.

7. Conclusion and Discussion

This study aims to qualitatively characterize the patterns of two participation networks in public and private policy-making and further empirically examines the presence of echo chambers in terms of information dissemination. Besides, the study provides a new network-based analytical framework focusing on the behavioral logic in the exploration and the collaborative interactions among multi-stakeholders related to transport policies. To summarize, the main conclusions are remarks from the findings that are listed as follows:

- (i) The descriptive analysis results provide evidence that multiple actors, especially the general public, share information in GCPN and ECPN. However, the participation networks are sparse and affected by a small component of participants with greater cyber influence and do not cover all crucial roles than others in information diffusion.
- (ii) The results suggest that when characterizing the key actors in information diffusion of both networks, organizations/companies and influential individuals (i.e., media companies, and journalists) contribute more actively as the “network root” to be at the central position, whereas the general public coming from various occupational backgrounds are more prominent as the “network bridge” to transfer the information further.
- (iii) Through the k-core decomposition method, the two studied networks exhibit the core-periphery structure as a small number of participants build a community with similar views but different backgrounds.
- (iv) The opinion analysis result reveals that the majority of participants will not express emotional viewpoints in transport policies. Furthermore, there is a general tendency that people are more likely to hold negative attitudes on social media.
- (v) The empirical results demonstrate the existence of echo chambers in both GCPN and ECPN, although its prevalence proportion is relatively low. Besides, echo chambers are configured around participants with critical opinions related to transport policies, and they generally happened such that organizations or influential individuals echoed the views of the general public.

These findings also have several managerial and policy implications. First, this study demonstrates the potential of using social media as a channel to investigate public participation among multi-stakeholders on major transport issues. Various measurements and indicators (e.g., network centrality, core-periphery) from the network perspective can be applied to analyze the relationship among actors. Second, it emphasizes the possibility of confirmation bias in transport policy networks that might be caused by the multi-stakeholders with diverse interests and their complex interactions. Policymakers and service providers should collect

a more complete sample of people’s views or consider diverse platforms in public engagement. Third, once the impact of confirmation bias is confirmed, it is suggested to conduct a further exploration based on the content of collected samples to improve the reliability of the result, such as topic models and sentiment analysis.

As a presentative of emerging mobility services, the service providers of bike-sharing are driven by complex interests, including being more likely to obtain a large number of user traffic instead of promoting travel services in the short term, gaining a higher market share to increase their valuation in the risk market, etc. Besides, its major users have certain characteristics, including being younger and willing to accept emerging techniques. Therefore, when using social media as a platform of public participation, bike-sharing policies may have a different/biased communication and participation process which is different from other urban transport policies. With the emergence of more new mobility services, the stakeholders may be more motivated and capable of intervening in transport policies. This is a topic that deserved continuous attention in the effective collaborative governance of urban transport. Empirical studies on a wider range of cases are needed to obtain a more profound understanding.

Despite its contributions and practical implications, this study still has several limitations. Its analysis is restricted to a specific participation network for bike-sharing price strategies, which may limit the generalizability of the empirical results to urban transport, other policy, and cultural contexts (e.g., different affecting or affected interest groups, the degree of willingness to express opinions through Internet, and participation channels available to the public). Further cross-disciplinary and cross-cultural research studies could be conducted to validate or expand the findings presented here. Furthermore, limited by the focus of revealing the key pattern of participation networks, this study only analyzes the forwarding/forwarded relationship represented by the single-directed, unweighted network. Further studies could include complete interaction relationships and features among participants included in the analysis.

Data Availability

The Weibo data used to support the findings of this study have not been made available because of the privacy protection rules of users’ data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors acknowledge support from the project of the National Natural Science Foundation of China (Nos. 71734004 and 71603186), Shanghai Pujiang Program (No. 2019PJC109), and China Scholarship Council (No. 201906260112). It was also partially supported by C2SMART

Tier 1 University Transportation Center at New York University.

References

- [1] L. Zhou, J. Ye, and X. Chen, *Current Situation and Problem Analysis of Public Participation in China's Urban Transportation Planning*, Transportation Research Board, Washington, DC, USA, 2015.
- [2] J. Gonzalez-Feliu, C. Pronello, and J. M. S. Grau, "Multi-stakeholder collaboration in urban transport: state-of-the-art and research opportunities," *Transport*, vol. 33, no. 4, pp. 1079–1094, 2018.
- [3] A. Alfayez, Z. Awwad, C. Kerr et al., "Understanding gendered spaces using social media data," in *Social Computing and Social Media: Applications and Analytics, Scsm 2017, Pt II*, G. Meiselwitz, Ed., vol. 10283, pp. 338–356, Springer, Cham, Switzerland, 2017.
- [4] L. Mandarano, M. Meenar, and C. Steins, "Building social capital in the digital age of civic engagement," *Journal of Planning Literature*, vol. 25, no. 2, pp. 123–135, 2010.
- [5] A. Gal-Tzur, S. M. Grant-Muller, E. Minkov, and S. Nocera, "The impact of social media usage on transport policy: issues, challenges and recommendations," Edited by J. F. DeSousa, J. P. DeSousa, A. Costa, T. Farias, and S. Melo, Eds., in *Transportation: Can We Do More with Less Resources?—16th Meeting of the Euro Working Group on Transportation—Porto 2013*, vol. 111, pp. 937–946, 2014.
- [6] S. Plous, *The Psychology of Judgment and Decision Making*, McGraw-Hill Book Company, New York, NY, USA, 1993.
- [7] E. Pariser, *The Filter Bubble: How the New Personalized Web Is Changing what We Read and How We Think*, Penguin, London, UK, 2011.
- [8] M. Salerno, T. Sanchez, S. Tomasello, and T. Metz, *Practices for Online Public Involvement (Project 20-05, Topic 49-11)*, Transportation Research Board, Washington, DC, USA, 2019, <http://www.trb.org/Publications/Blurbs/179304.aspx>.
- [9] A. Gal-Tzur, S. M. Grant-Muller, T. Kuflik, E. Minkov, S. Nocera, and I. Shoor, "The potential of social media in delivering transport policy goals," *Transport Policy*, vol. 32, pp. 115–123, 2014.
- [10] Y. Cheng, "Collaborative planning in the network: consensus seeking in urban planning issues on the Internet—the case of China," *Planning Theory*, vol. 12, no. 4, pp. 351–368, 2013.
- [11] S. M. Kaufman, *How Social Media Moves New York: Twitter Use by Transportation Providers in the New York Region*, Transportation Research Board, Washington, DC, USA, 2012, http://wagner.nyu.edu/files/faculty/publications/how_social_media_moves_new_york.pdf.
- [12] B. Batrinca and P. C. Treleaven, "Social media analytics: a survey of techniques, tools and platforms," *Ai & Society*, vol. 30, no. 1, pp. 89–116, 2015.
- [13] C. Collins, S. Hasan, and S. V. Ukkusuri, "A novel transit rider satisfaction metric: rider sentiments measured from online social media data," *Journal of Public Transportation*, vol. 16, no. 2, p. 2, 2013.
- [14] L. Tang, Z. Duan, and Y. Zhao, "Toward using social media to support ridesharing services: challenges and opportunities," *Transportation Planning and Technology*, vol. 42, no. 4, pp. 355–379, 2019.
- [15] M. Zhao, Y. Lin, and B. Derudder, "Demonstration of public participation and communication through social media in the network society within Shanghai," *Environment and Planning B: Urban Analytics and City Science*, vol. 45, no. 3, pp. 529–547, 2017.
- [16] B. A. Furtado, M. A. Fuentes, and C. J. Tessone, "Policy modeling and applications: state-of-the-art and perspectives," *Complexity*, vol. 2019, Article ID 5041681, 11 pages, 2019.
- [17] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Vol. 8, Cambridge University Press, Cambridge, UK, 1994.
- [18] J. Scott, *Social Network Analysis: A Handbook*, Sage publications, London, UK, 2nd edition, 2000.
- [19] L. Xie, T. Han, and M. Skitmore, "Governance of relationship risks in megaprojects: a social network analysis," *Advances in Civil Engineering*, vol. 2019, Article ID 1426139, 13 pages, 2019.
- [20] Y. Tang, G. Wang, H. Li, D. Cao, and X. Li, "Comparing project-based collaborative networks for BIM implementation in public and private sectors: a longitudinal study in Hong Kong," *Advances in Civil Engineering*, vol. 2019, Article ID 6213694, 14 pages, 2019.
- [21] D. P. Cheung and M. H. Gunes, "A complex network analysis of the United States air transportation," in *Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Istanbul, Turkey, August 2012.
- [22] M. G. Song and G. T. Yeo, "Analysis of the air transport network characteristics of major airports," *The Asian Journal of Shipping and Logistics*, vol. 33, no. 3, pp. 117–125, 2017.
- [23] S. Dörry and A. Decoville, "Governance and transportation policy networks in the cross-border metropolitan region of Luxembourg: a social network analysis," *European Urban and Regional Studies*, vol. 23, no. 1, pp. 69–85, 2013.
- [24] H. Sun, Y. Zhang, Y. Wang, L. Li, and Y. Sheng, "A social stakeholder support assessment of low-carbon transport policy based on multi-actor multi-criteria analysis: the case of Tianjin," *Transport Policy*, vol. 41, pp. 103–116, 2015.
- [25] K. Weiss, M. Hamann, M. Kinney, and H. Marsh, "Knowledge exchange and policy influence in a marine resource governance network," *Global Environmental Change*, vol. 22, no. 1, pp. 178–188, 2012.
- [26] M. Mills, J. G. Álvarez-Romero, K. Vance-Borland et al., "Linking regional planning and local action: towards using social network analysis in systematic conservation planning," *Biological Conservation*, vol. 169, pp. 6–13, 2014.
- [27] R. S. Nickerson, "Confirmation bias: a ubiquitous phenomenon in many guises," *Review of General Psychology*, vol. 2, no. 2, pp. 175–220, 1998.
- [28] P. M. Wagner and T. Ylä-Anttila, "Can policy forums overcome echo chamber effects by enabling policy learning? Evidence from the Irish climate change policy network," *Journal of Public Policy*, vol. 40, no. 2, pp. 194–211, 2018.
- [29] K. Garimella, G. De Francisci Morales, A. Gionis, and M. Mathioudakis, "The effect of collective attention on controversial debates on social media," in *Proceedings of the 2017 ACM*, pp. 43–52, Troy, MI, USA, June 2017.
- [30] M. Del Vicario, G. Vivaldo, A. Bessi et al., "Echo chambers: emotional contagion and group polarization on Facebook," *Scientific Reports*, vol. 6, no. 1, p. 37825, 2016.
- [31] M. Cinelli, G. D. F. Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "Echo chambers on social media: a comparative analysis," 2020, <https://arxiv.org/abs/2004.09603>.
- [32] A. Bessi, M. Coletto, G. A. Davidescu, A. Scala, G. Caldarelli, and W. Quattrociocchi, "Science vs conspiracy: collective narratives in the age of misinformation," *PLoS One*, vol. 10, no. 2, Article ID e0118093, 2015.
- [33] A. Cossard, G. D. F. Morales, K. Kalimeri et al., "Falling into the echo chamber: the Italian vaccination debate on Twitter," 2020, <https://arxiv.org/abs/2003.11906>.

- [34] C. R. Sustain, *Republic. Com 2.0*, Princeton University Press, Princeton, NJ, USA, 2009.
- [35] “Facebook’s CEO Mark Zuckerberg’s open letter,” 2017, <https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471/>.
- [36] L. Jasny, J. Waggle, and D. R. Fisher, “An empirical examination of echo chambers in US climate policy networks,” *Nature Climate Change*, vol. 5, no. 8, pp. 782–786, 2015.
- [37] E. Colleoni, A. Rozza, and A. Arvidsson, “Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data,” *Journal of Communication*, vol. 64, no. 2, pp. 317–332, 2014.
- [38] L. Jasny, A. M. Dewey, A. G. Robertson et al., “Shifting echo chambers in US climate policy networks,” *PLoS One*, vol. 13, no. 9, Article ID e0203463, 2018.
- [39] F. Ma, W. Shi, K. F. Yuen, Q. Sun, and Y. Guo, “Multi-stakeholders’ assessment of bike sharing service quality based on DEMATEL-VIKOR method,” *International Journal of Logistics Research and Applications*, vol. 22, no. 5, pp. 449–472, 2019.
- [40] *Ofo Faces Flak on Refunds of Deposits*, China Daily, Beijing, China, 2018, <https://www.chinadaily.com.cn/a/201812/20/WS5c1ad088a3107d4c3a001bde.html>.
- [41] *Tighter Controls on Transport Deposits*, China Daily, 2019, <http://global.chinadaily.com.cn/a/201905/17/WS5cde0cdfa3104842260bc30b.html>.
- [42] C. Noy, “Sampling knowledge: the hermeneutics of snowball sampling in qualitative research,” *International Journal of Social Research Methodology*, vol. 11, no. 4, pp. 327–344, 2008.
- [43] R. A. Hanneman and M. Riddle, “A brief introduction to analyzing social network data,” *The Sage Handbook of Social Network Analysis*, pp. 331–339, SAGE Publications Ltd, Thousand Oaks, CA, USA, 2011.
- [44] A. Punel and A. Ermagun, “Using Twitter network to detect market segments in the airline industry,” *Journal of Air Transport Management*, vol. 73, pp. 67–76, 2018.
- [45] M. Newman, “Power laws, Pareto distributions and Zipf’s law,” *Contemporary Physics*, vol. 46, no. 5, pp. 323–351, 2005.
- [46] L. C. Freeman, “Centrality in social networks conceptual clarification,” *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.
- [47] L. C. Freeman, “A set of measures of centrality based on betweenness,” *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
- [48] I. Alvarez-Hamelin, L. Dall’Asta, A. Barrat, and A. Vespignani, “k-core decomposition: a tool for the visualization of large scale networks,” 2005, <https://arxiv.org/abs/cs/0504107>.
- [49] S. P. Borgatti, M. G. Everett, and L. C. Freeman, *Ucinet for Windows: Software for Social Network Analysis*, Analytic Technologies, Harvard, MA, USA, 2002.
- [50] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [51] M. Castells, “Network theory—A network theory of power,” *International Journal of Communication*, vol. 5, p. 15, 2011.
- [52] L. L. Berry, “The collaborative organization:,” *Organizational Dynamics*, vol. 33, no. 3, pp. 228–242, 2004.

Research Article

Predicting Critical Bicycle-Vehicle Conflicts at Signalized Intersections

Alireza Darzian Rostami , **Anagha Katthe** , **Aryan Sohrabi**, and **Arash Jahangiri** 

Department of Civil, Construction, and Environmental Engineering, San Diego State University, 5500 Campanile Dr, San Diego, CA 92182, USA

Correspondence should be addressed to Arash Jahangiri; ajahangiri@sdsu.edu

Received 29 August 2020; Revised 17 October 2020; Accepted 16 November 2020; Published 3 December 2020

Academic Editor: Kun Xie

Copyright © 2020 Alireza Darzian Rostami et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Continuous development of urban infrastructure with a focus on sustainable transportation has led to a proliferation of vulnerable road users (VRUs), such as bicyclists and pedestrians, at intersections. Intersection safety evaluation has primarily relied on historical crash data. However, due to several limitations, including rarity, unpredictability, and irregularity of crash occurrences, quantitative and qualitative analyses of crashes may not be accurate. To transcend these limitations, intersection safety can be proactively evaluated by quantifying near-crashes using alternative measures known as surrogate safety measures (SSMs). This study focuses on developing models to predict critical near-crashes between vehicles and bicycles at intersections based on SSMs and kinematic data. Video data from ten signalized intersections in the city of San Diego were employed to train logistic regression (LR), support vector machine (SVM), and random forest (RF) models. A variation of time-to-collision called T_2 and post-encroachment time (PET) were used to specify monitoring periods and to identify critical near-crashes, respectively. Four scenarios were created using two thresholds of 5 and 3 s for both PET and T_2 . In each scenario, five monitoring period lengths were examined. The RF model was superior compared to other models in all different scenarios and across different monitoring period lengths. The results also showed a small trade-off between model performance and monitoring period length, identifying models with monitoring period lengths of 10 and 20 frames performed slightly better than those with lower or higher lengths. Sequential backward and forward feature selection methods were also applied that enhanced model performance. The best RF model had recall values of 85% or higher across all scenarios. Also, RF prediction models performed better when considering just the rear-end near-crashes with recalls of above 90%.

1. Introduction

With a growing interest in using eco-friendly modes of travel such as bicycling and walking, there is an increasing trend in the number of crashes involving. Comparing to vehicular modes of transportation, bicyclists and pedestrians are at an increased risk of fatal or severe injuries at the time of the collision. Based on the National Highway Traffic Safety Administration (NHTSA) report [1] in 2018, 857 bicycle fatalities occurred in the US roadways, which was 6% more than the preceding year and the highest number since 1990. Bicyclist fatalities had increased from 1.9% of total road crash fatalities in 2009 to 2.3% in 2018. Intersections are considered as hot spots when looking at

traffic fatalities and injuries due to the complex nature of interactions. The presence of mixed traffic flow specifically makes intersections more important resulting in a large number of traffic incidents and collisions. According to the Federal Highway Administration (FHWA), an average of 2.5 million crashes occur every year at intersections. Traffic Safety Facts published by NHTSA in 2018 indicates that 29% of pedalcyclist fatalities happened at intersections. Road-user movements possessing distinct conflicting patterns could lead to a large number of crashes at signalized intersections [2]. Therefore, it is crucial to understand the dynamics of interactions between bicyclists and motor vehicles by looking at their critical conflicts.

In general, direct or indirect techniques can be used to perform safety analysis [3]. Direct techniques hinge on historical crash data to evaluate the degree of road safety. Traditionally, crash frequency and crash severity have been used as metrics for safety evaluations. Crash data help in identifying high-risk locations, current conditions, necessary safety improvements, and safety evaluation of road users [4, 5]. However, unpredictability and irregularity of crash occurrences in the real environment result in the inefficiency of quantitative and qualitative crash determination. Due to the possibility of erroneous or inconsistent reporting, crash data are unreliable [6]. When it comes to improving intersection safety, crash data have several limitations. First, as roadway crashes are rare events, it takes a long period of time for meaningful crash data to be available. This is more so the case when a specific crash type is being studied (e.g., a crash between bicyclists making left turns from an approach and vehicles going through the intersection from the opposite approach). Moreover, changes such as design improvements and demand increase could occur during such long periods, potentially impacting safety evaluation outcomes. Second, reporting of crash events largely depends on the type of road users involved and the severity of crash events and is also found to be unevenly distributed [7]. Specifically, vulnerable road users are heavily underreported, thus making it unreliable and less accurate to use crash data for safety evaluation. Third, although crash data specify the occurrence of the event, it has a limited scope of information regarding the cause of the crash that includes precrash movements, distinct road-user behavior patterns, and other situational aspects [7]. Last, crash data analysis is a reactive approach in which remedial measures can be incorporated only after the occurrence of crashes, and hence, critical locations are identified after observing multiple fatalities and injuries, and countermeasures are implemented after the fact [8].

Given these shortcomings, indirect safety indicators have been studied [7, 9–11]. Traffic conflict technique (TCT) was proposed [12, 13] that utilizes conflicts (i.e., near-crashes) as a substitute for actual crashes [14, 15]. A traffic conflict is an “observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged” [16]. TCT concept classifies various traffic events based on conflict severity level and represents road-user interactions as a continuum of safety-related events. These safety-related events range from slight conflicts, potential conflicts, and severe conflicts [17, 18]. The severity and frequency of these conflicts are typically quantified by measures known as surrogate safety measures (SSMs). Critical near-crashes are detected if an SSM value crosses a critical threshold [19].

Frequency estimation of near-crashes using SSMs has been conducted in the past and still is ongoing research. Various SSMs have been used, and among them, time-to-collision (TTC) and postencroachment time (PET) have been widely used to evaluate safety at signalized intersections. At an instant t , TTC is defined as the “time taken by the two road users to collide, provided the collision course

and speed difference are the same” [20]. Considering the trajectories of two road users, the relative time-to-collision (RTTC) is estimated as the difference in time taken by the first and second road users to reach a predicted trajectory intersection point (TIP). TTC is considered as a special case of RTTC where the RTTC value is zero (i.e., both users reach TIP at the same time resulting in a collision). As RTTC decreases, the conflict is considered more critical. However, a low value of RTTC could also arise from cases when two objects that are far away from the TIP have close arrival times to the TIP. Therefore, RTTC alone could not be a good indicator of critical near-crashes. To address this issue, Laureshyn et al. [7] proposed T_2 as the predicted time taken by the latest road user to reach the conflict points. A low value of T_2 indicates time remained for the second road user to perform braking or do an evasive maneuver to prevent or mitigate potential collisions. PET is defined as the “time between the departure of the encroaching vehicle from the conflict point and the arrival of the vehicle with the right of way at the conflict point” [14, 21]. Unlike TTC and RTTC, PET provides a single value for every road-user pair crossing the path of each other. As PET is calculated based on an observed TIP (as opposed to “predicted TIP” in the TTC case), it is considered as a tangible indicator of how critical an interaction is based on what actually occurred as opposed to what was predicted.

Since SSMs can be used to identify unsafe events proactively, it is of interest to develop models to predict whether an interaction is going to be safe. Timely prediction of near-crashes could be used in countermeasures such as changing the signal settings to avoid or mitigate crashes or warning approaching road users about potential collisions in a connected environment. However, there are a limited number of studies on this topic, especially for vehicle-bicycle interactions. The goal of this research is to develop models for bicyclists at signalized intersections to predict critical bicycle-vehicle near-crashes. Kinematic features such as velocity and acceleration as well as the SSMs were used in the model development process. The prediction models were developed using logistic regression (LR), support vector machine (SVM), and random forest (RF) methods. The remainder of the paper is organized as follows: past studies are reviewed in the literature review section. Next, data preparation tasks are described. Subsequently, scenario development using different SSMs is explained followed by the model development section. Finally, the model results, conclusions, and potential future research are discussed.

2. Literature Review

Modeling crash probability and severity has long been studied among researchers and practitioners. Conventional safety performance functions (SPFs) estimate the crash frequency of a road segment or intersection as a function of explanatory variables describing the characteristics of roadway design and environment, vehicles, and humans [22, 23]. Due to the rare and sporadic nature of crashes, statistical count models such as Poisson and negative binomial (NB) and their variants are frequently used for

analyzing explanatory variables and predicting crash frequency [24–26]. Besides statistical count models, artificial intelligence models have also been proposed for crash frequency prediction [27–29]. Typical SPFs consider design features such as speed limits and intersection type as well as aggregated measures like annual average daily traffic (AADT) to identify hazardous locations and missing information about individual driver behavior. However, it was found that variations in traffic dynamics across individuals could be a potential contributing factor to the crashes, which can be utilized to predict the occurrence of crashes in real time [22]. Explanatory variables related to traffic dynamics include speed, flow, and occupancy obtained from real-time traffic data [30].

Several studies were conducted considering traffic dynamics data along with crash data to predict crash likelihood at different road types and intersections in real time. Oh et al. [30] defined a 5-minute period right before a crash as a disruptive traffic condition and a 5-minute period 30 minutes before the crash as a normal traffic condition. Using the two traffic conditions, an SPF model was developed using the nonparametric Bayesian approach to assess the likelihood of crashes in real time. It was found that the standard deviation of speed was the most reliable indicator to develop an accurate SPF. In the real-time application of this model, traffic dynamics were continuously monitored in 5-minute intervals to estimate the crash likelihood after each interval. If the likelihood of crash was above a certain threshold, the driver would be alerted to either reduce/increase the speed to minimize the overall speed variation. Similarly, Hossain and Muromachi [31] developed a Bayesian network model using traffic data collected from an expressway over 16 months to predict crashes in real time. This study considered 5-minute aggregated average speed and cumulative flow as the main predictive variables. Five minutes before a crash termed as “precrash traffic conditions” and corresponding average speed and cumulative flow were used to predict the probability of crashes. To improve the performance of the SPF, a hybrid model combining support vector machine and *k*-means clustering algorithm was developed that promised better crash predictions [32]. Traffic flow data for 5–10 minutes before the occurrence of a crash were considered in this study. Speed, volume, and traffic flow difference between upstream and downstream and average traffic flow were the predictive variables considered in this study. To extend the analysis to signalized intersections, Yuan and Abdel-Aty [33] developed a Bayesian conditional logistic model to predict real-time crashes within the intersection and also at the entrance of the intersection. This model obtained the best results when traffic data were taken from the 5- to 10-minute time period before the crash. In addition to real-time traffic data, signal timing and weather conditions were considered in order to predict crash likelihood.

Most crash prediction models utilize historical crash data that have some limitations, as described earlier. In order to overcome these limitations, crash surrogates, also called “near-crashes” or “traffic conflicts,” have been proposed to see if they can serve as reliable indicators of safety. El-Basyouny and Sayed [34] developed SPFs based on traffic

conflicts at signalized intersections and found a significant proportional relationship between traffic conflicts and crashes. In a similar study, Sacchi and Sayed [35] developed SPFs from traffic conflicts for each crash type. Some efforts [36–38] focused on developing crash probabilistic frameworks that estimate collision risk based on road users’ hypothetical future movements and the traffic conflict concept. For example, Fu et al. [36] used a distance-velocity (DV) framework to study secondary pedestrian-vehicle interactions at nonsignalized intersections. The framework used in this study incorporates road-user kinematic and SSMs data with driver reaction time to assess the safety of interactions between road users. Several studies investigated the relationship between crashes and near-crashes. Parker and Zegeer [39] showed a statistically significant linear relationship between crashes and near-crashes. He et al. [40] used the safety pilot model deployment (SPMD) to study the statistical relationship between SSMs and crashes and found that modified time-to-collision (MTTC) performed better than deceleration to avoid a crash (DRAC) and TTC.

Many studies proved the benefits of the extreme value theory approach that estimates the risk of crashes (extreme events) based on the risky interactions between road users [41–44]. In a recent study, this approach was used to predict crashes at a signalized intersection based on traffic conflicts extracted from microscopic traffic simulation and real-world data [45]. The result showed a high correlation between simulated and actual field traffic conflict and a promising result for crash estimation using field conflict. Zheng and Sayed [46] compared four SSMs, including TTC, PET, modified time-to-collision (MTTC), and deceleration to avoid a crash (DRAC), to estimate crashes at signalized intersections and found that the MTTC produced the most accurate result. Guo et al. [47] evaluated the use of near-crashes to assess the safety and showed a positive correlation between causes for crashes and near-crashes. Similar analyses were conducted by [48, 49] confirming a positive correlation between crashes and near-crashes.

Although SSMs have been used in numerous studies to evaluate the safety or investigate their correlations to actual crashes, there is limited research on using SSMs to predict near-crashes, as summarized in Table 1. Some studies used regression analysis to estimate the frequency of near-crashes over a time period. For example, Essa and Sayed [52, 54] developed a safety performance model to predict traffic conflicts in each cycle at signalized intersections using a set of explanatory variables such as traffic volume, queue length, and platoon ratio. Ma et al. [53] developed a model to predict an hourly conflict risk index (HCRI) for expressway diverging areas using variables such as traffic volume and speed in the mainline and the ramp areas. HCRI was calculated based on rear-end and lane changing conflicts identified by TTC.

Other efforts have been done on predicting and classifying the upcoming near-crashes between a pair of road users in real time. Utilizing near-crash data from SHRP2 NDS datasets, Osman et al. [51] applied several supervised machine learning models to predict rear-end near-crashes a few seconds before they happen. The standard deviations of

TABLE 1: Near-crash prediction studies.

Study	Conflict type	Interaction type	Methods	Identifying near-crashes	Location	Independent variable	Evaluation metric
Formosa et al. [50]	Rear-end, lane changing	Vehicles-vehicles	Deep learning (DP)	Time headway, lateral acceleration, lateral distance, spacing, speed, etc.	Roadway segment	26 variables including SSMs	Classification: accuracy, AUC, precision, recall, and false alarm ratio
Osman et al. [51]	Rear-end	Vehicles-vehicles	KNN, random forest, SVM, decision trees, Gaussian NB, and AdaBoost	Defined in the SHRP2 NDS database	Roadway segment	Std of kinematic data	Classification: accuracy, recall, precision, and F1
Essa and Sayed [52]	Rear-end	Vehicles-vehicles	The safety performance functions (SPFs) were developed using the full Bayesian approach	TTC, MTTTC, and DRCA	Signalized intersections	Traffic volume, shock wave area, maximum queue length, backward-moving shock wave speed, platoon ratio	Regression: goodness of fit, the scaled deviance (SD), and the Pearson chi-squared (χ^2)
Ma et al. [53]	Rear-end, lane changing	Vehicles-vehicles	Multivariate linear regression	TTC	Expressway diverging areas and ramps	Traffic volume, speed, etc.	Regression: goodness of fit R^2

the vehicle kinematics data such as acceleration, yaw rate, speed, and pedal position during a monitoring period before the conflicts were used as independent variables to identify the upcoming unsafe events. The time before a traffic conflict was divided into two periods: the turbulence horizon, which is the time to monitor and record the kinematic data, followed by the prediction period, which is the time for the model to predict the outcome before the conflict. Results from sensitivity analysis showed 1 second as optimal prediction horizon length and 3 seconds as turbulence horizon. This study concluded that near-crash prediction models are highly efficient in predicting most instances of near-crashes with minimum false predictions. Formosa et al. [50] combined highly disaggregated traffic data taken from motorway incident detection and automatic signaling (MIDAS) with various SSMs including TTC and PET calculated from in-vehicle sensors data from an instrumented vehicle using centralized, integrated data architecture to develop a real-time traffic conflict detection and prediction method using deep learning methodology. A total of 26 input variables were used consisting of six widely used SSMs, the mean and standard deviation of ego vehicle's speed and speed variance between lanes and traffic variables by lane such as speed, density, flow, headway, and occupancy. The safe traffic dynamics and traffic conflicts were identified using time headways between the probe vehicle and the leading vehicle for rear-end and lane changing conflicts. It was found that TTC varied by speed, weather, and traffic density. The best deep neural network model provided an accuracy of 94%. However, this study is limited to lane change and rear-end type of conflicts.

2.1. Identification of Critical Events. Although SSMs provide a means to quantify near-crash events, there is a necessity to distinguish such events as critical or noncritical by considering a threshold value for the SSM of interest. Studies on predicting critical near-crashes took different approaches to distinguish between critical and noncritical near-crashes. Formosa et al. [50] considered several criteria for time headway, lateral distance, and speed of following and leading vehicle to identify critical lane changing and rear-end conflict. Essa and Sayed [52] used a range of threshold values for TTC, modified time-to-collision (MTTC), and deceleration rate to avoid a crash (DRAC) to label near-crashes, and they reported the result for all of these variations. Ma et al. [53] recruited a group of students to watch the video and find critical near-crashes. Then, they used a 85-percentile of TTC value as the threshold for this SSM. There has not been a consensus among researchers to use consistent threshold values. This could be due to various factors such as type of the road, different driver behaviors in different locations, type of conflicts, and type of SSMs used [55, 56]. Mahmud et al. [57] provided a review of 38 common SSMs and their thresholds. To find the optimal SSMs threshold, a recent study [58] investigated a range of thresholds for SSMs and chose the ones that maximize the correlation coefficient between crash data and risk associated with SSMs. Sayed et al. [59] included only bicycle-vehicle interactions that had

TTC value less than 3 s in their safety evaluation. In another study, the TTC and PET threshold values were considered as 1.5 seconds for interactions between vehicles (including taxis), lorries (including bus), pedestrians, and bicyclists [60]. Considering vehicle-bicycle interactions, Zangenehpour et al. [61] used both TTC and PET threshold values less than 5 seconds and less than 1.5 seconds for labeling them as conflicts and dangerous conflicts, respectively. In some studies [62, 63], conflicts between vehicles and bicycles were classified into more than two categories; conflicts were categorized as very dangerous interactions for $PET \leq 1.5$ seconds, dangerous interactions for $1.5 \text{ seconds} < PET < 3$ seconds, mild interactions for $3.0 \text{ seconds} < PET \leq 5.0$ seconds, and no interactions for $PET > 5$ seconds. Table 2 summarizes some of the common SSMs threshold used in previous studies.

3. Materials and Methods

This study aims to develop a model to predict critical bicycle-vehicle interactions at signalized intersections. The model response variable is a dichotomous variable (critical or noncritical) that was labeled using the PET value. In addition, this study utilizes T_2 (a variation of time to the collision as described earlier) to define a monitoring period during which road-user kinematic data were extracted for prediction model development. The methodology of this study is summarized in Figure 1.

3.1. Data Collection and Preparation. In this study, we used video data collected from cameras installed at ten signalized intersections in the city of San Diego. The video data were collected for a period of 24 hours for each intersection on a workday (Tuesday, Wednesday, or Thursday) in the month of May, June, or July 2018. The original data were collected as part of a previous study [65]. The ten signalized intersections utilized in the present study were chosen based on their higher bicycle activities identified by manually reviewing video recordings. The ten data collection sites are shown in Figure 2 and Table 3.

The video data were then reviewed focusing on morning, afternoon, and evening peak hours. Based on manual observation of the maximum activity of bicyclists, video data were filtered further down to 5-minute clips to focus on bicycle-vehicle interactions. A video data analysis was performed on raw video data in order to extract the road-users' trajectory and transform it into the top-down view (bird's-eye view). The detailed procedure of this process can be found in [65], which discusses tasks such as data annotation, object detection, and trajectory extraction. Figure 3 shows an example of the annotation task that is essential in identifying object trajectories. While the machine vision modeling was not the focus of the present study, the outputs from the machine vision models were used including the road-users type and the location at each frame. Each frame in this study is 1/30 of a second.

Vehicles and bicyclists were monitored as they interacted with each other to calculate SSMs and kinematic features for

TABLE 2: PET and TTC thresholds used in previous studies.

SSMs	Threshold under	Studies
PET	1.5	[61–63]
	3	[62–64]
	5	[61–64]
TTC	1.5	[60]
	3	[45, 59]
	5	[61]

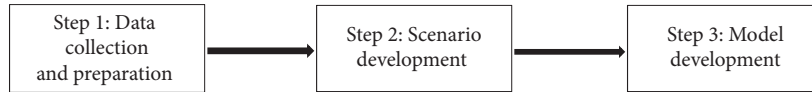


FIGURE 1: Methodology framework.

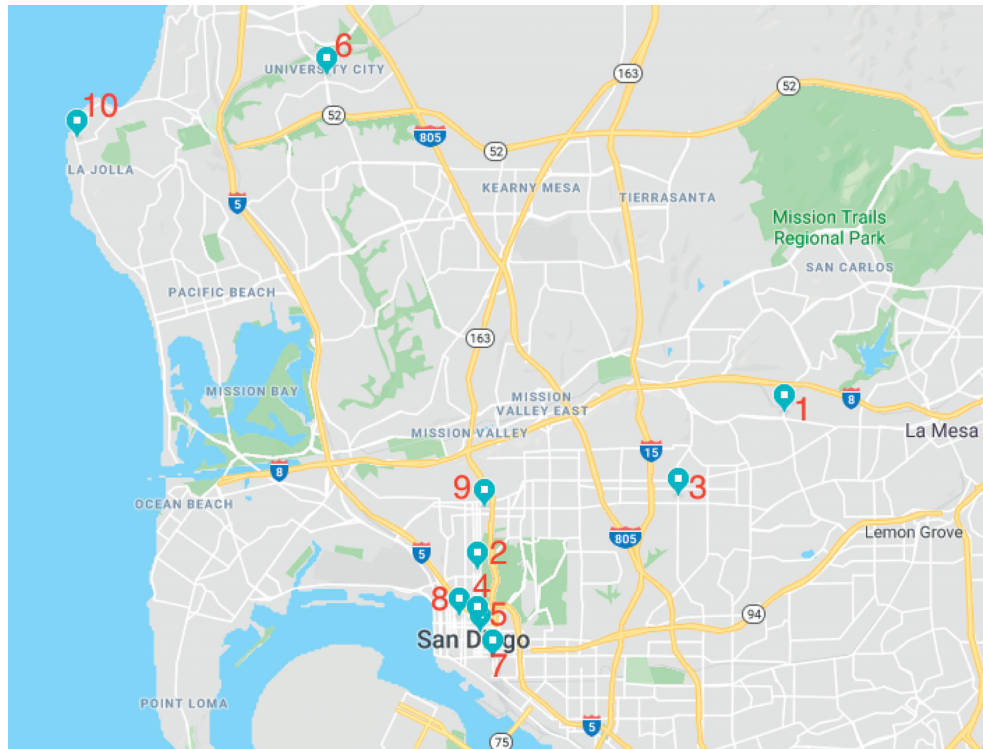


FIGURE 2: Study intersections.

TABLE 3: Study intersections.

Intersection number	Name of the intersection
1	College Ave & Montezuma Rd
2	5th Ave & Laurel St
3	Fairmount Ave & University Ave
4	Fifth Ave & B St
5	Sixth Ave & Broadway
6	Genesee Ave & Governor Dr
7	10th Ave & J St
8	Union St & Ash St
9	7th Ave & Robinson Ave
10	La Jolla Blvd & Pearl St

each interaction. Several kinematic features including velocity, acceleration, relative approach distance, and relative approach velocity were computed for both vehicles and bicycles at each frame. For each vehicle and bicycle that interacted with each other, the TIPs were predicted utilizing the velocities and headings of the two objects. Next, time taken by each of the objects to reach the predicted TIP was determined as the time to intersection (TTX_1 and TTX_2 , respectively). TTX and its derivatives are continuous variables, meaning that as long as the predicted path of two road-users crosses each other, these values could be calculated. Furthermore, PET was estimated by calculating the difference between the time frame when the first road-user

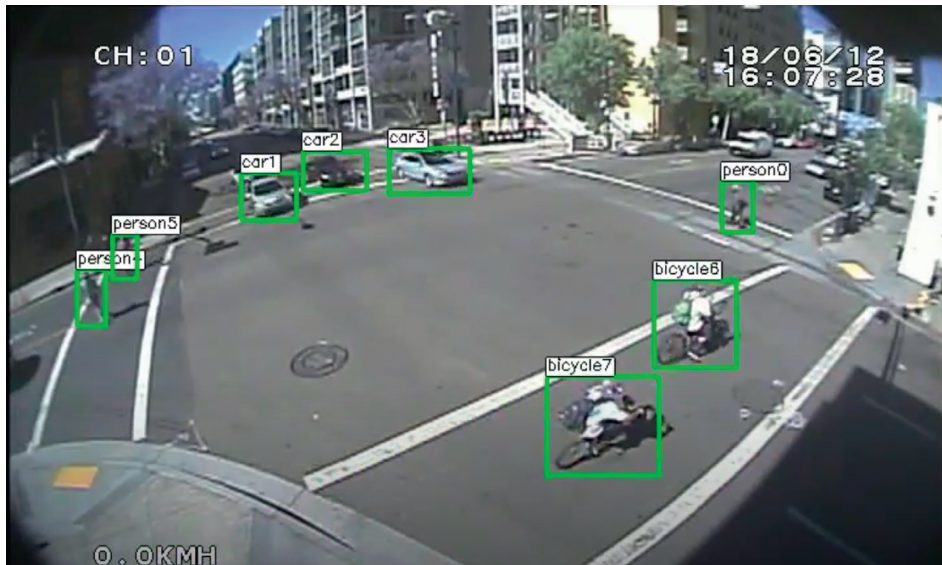


FIGURE 3: Example of annotation output.

departs from the observed TIP and the time frame when the second one reaches the same observed TIP. Table 4 demonstrates all features that were used in this study.

3.2. Scenario Development. Scenarios were developed using two parameters: (1) PET values were employed to identify “critical (C)” or “noncritical (NC)” near-crashes. Two commonly used values of 3 s and 5 s were used to identify critical near-crashes. A near-crash was labeled as critical when the PET value goes below the defined threshold. Also, near-crashes with no PET value or PET value higher than the threshold were labeled as noncritical. (2) As shown in Figure 4, T_2 values were used to define monitoring periods. The end of the monitoring period (T_E) was designated as the first time the T_2 value reaches below a defined threshold during an interaction. Two thresholds of 3 and 5 s for T_2 were utilized. The interactions that did not have a T_2 value below the threshold were not considered as a near-crash. For example, the T_2 value of 5 seconds means that in an interaction, the second road-user (farthest from the TIP) has 5 seconds to perform a maneuver to avoid a potential crash. Stepping 1, 5, 10, 20, and 30, frames backward from the monitoring period end points marked different starting points (T_S), resulting in five different lengths of monitoring periods (1/30 s, 1/6 s, 1/3 s, 2/3 s, and 1 s, respectively). As a result, four scenarios were developed for near-crash prediction models using different PET (i.e., 3 and 5) and T_2 (i.e., 3 and 5) thresholds. Each scenario was examined for five different monitoring periods. As illustrated in Figure 4, kinematic features and SSMs such as RTTC were recorded during the monitoring period until T_2 reaches below the threshold (e.g., 3) at T_E . Kinematic data recorded at each time frame during a monitoring period resulted in a series of values for each feature. Therefore, statistical measures such as mean and standard deviation were used to reduce multiple values into one. Due to sudden changes in speed and travel direction, the value of SSMs was not present at all

frames during an interaction. Therefore, the last value of these features recorded at the end of the monitoring period was used. The last recorded value of SSMs is also more indicative of the severity of an interaction as it presents the situation closest to the TIP. Table 5 shows all the 30 initial features used in this study.

3.3. Model Building and Analysis. Predicting bicycle-vehicle conflicts as “critical conflicts” and “noncritical conflicts” is a classification problem. Three popular supervised machine learning classifiers of logistic regression, SVM, and random forest were used to build vehicle-bicycle near-crash prediction models. Logistic regression is a commonly used machine learning classifier offering ease of implementation and interpretation and fast computation. Also, it has been extensively used in crash prediction literature and mostly as a baseline model. Since the dependent variable in this study is dichotomous (critical or noncritical near-crashes), we used binary logistic regression to describe the relationship between the dependent and independent variables. SVM is a supervised machine learning classifier that can be used to solve both classification and regression problems. This method plots each data point in the n -dimensional space defined by the input features. Then, the algorithms find the best hyperplane that could divide the classes with the largest gap between classes. Depending on the kernel function, SVM could perform linear and nonlinear classification boundaries. SVM is suitable for cases with high-dimensional input features, and it could perform well in facing the outliers and extreme cases in binary classification, leading to good generalization. Random forest (RF) is an ensemble learning method that can be used for both classification and regression problems. It generates several random decision trees at training time and relies on the majority of votes across these trees to predict classes for each observation. RF performs well on imbalanced datasets, and since it generates several trees on the subset of the data and combines the

TABLE 4: Prediction model features.

Features	Description
Velocity	Velocity computed for interacting vehicle and bicycle at every frame
Acceleration	Acceleration computed for interacting vehicle and bicycle at every frame
Relative distance	The Euclidean distance between interacting vehicle and bicycle in the same frame
Relative velocity	Rate of change of relative distance computed at every frame
TTX_{avg}	Mean of TTX_1 and TTX_2
RTTC	Absolute difference between TTX_1 and TTX_2
T_2	T_2 is defined as the maximum value of TTX_1 and TTX_2
Direction of conflict	1: if the angle between two road users is between 0° and 15° , and 0: otherwise 1 indicates a rear-end type of near-crash. Otherwise, it could be sideswipe, lane changing, or other types of near-crashes
Distance to conflict point	Distance from the current road-user location and the predicted TIP

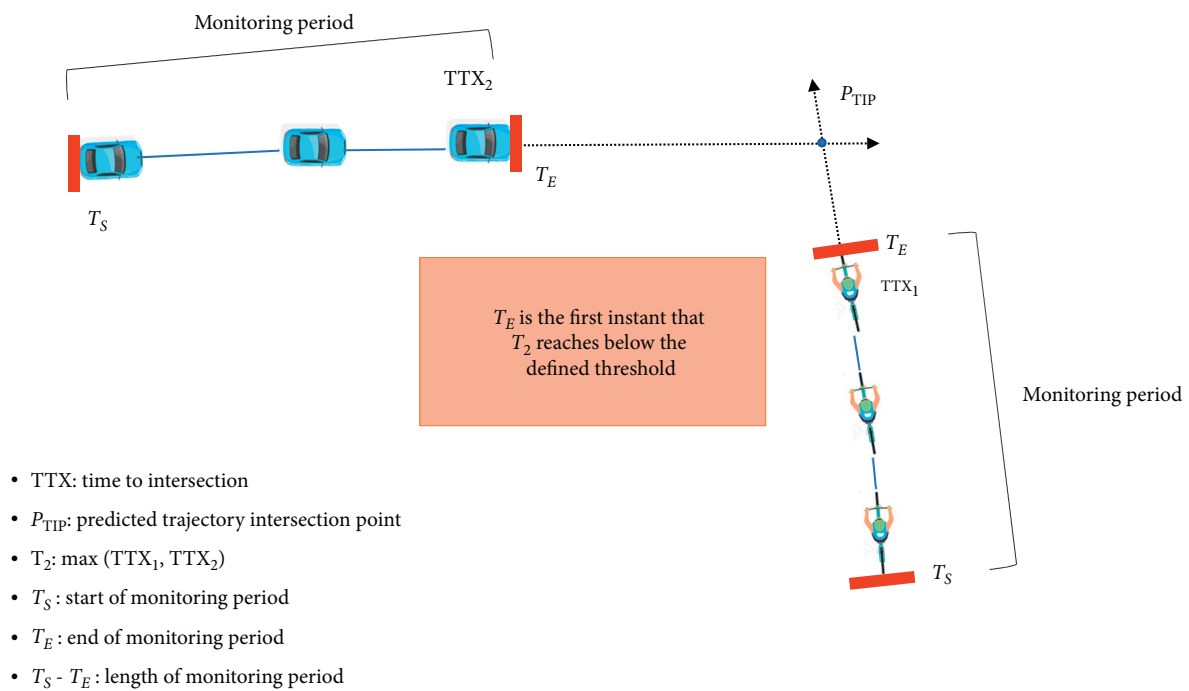


FIGURE 4: Scenario development parameters.

TABLE 5: Input features.

Kinematic features	(i) Mean of object 1 velocity, max of object 1 velocity, min of object 1 velocity, and the standard deviation of object 1 velocity
	(ii) Mean of object 2 velocity, max of object 2 velocity, min of object 2 velocity, and the standard deviation of object 2 velocity
	(iii) Mean of object 1 acceleration, max of object 1 acceleration, min of object 1 acceleration, and the standard deviation of object 1 acceleration
	(iv) Mean of object 2 acceleration, max of object 2 acceleration, min of object 2 acceleration, and the standard deviation of object 2 acceleration
	(v) Mean of relative distance, max of relative distance, min of relative distance, and the standard deviation of relative distance
	(vi) Mean of relative velocity, max of relative velocity, min of relative velocity, and the standard deviation of relative velocity
SSM-related features	T_2 , RTTC, TTX_{avg} , object 1 distance to conflict point, object 2 distance to conflict point, and direction of conflict

output of all the trees, it reduces overfitting and variance problems.

In this study, logistic regression and SVM were implemented using the Python scikit-learn library. The random forest classifier was implemented using the `BalancedRandomForestClassifier` in the Python `imbalanced-learn` package. In this implementation of random forest, random undersampling of the majority class (noncritical near-crashes) was performed on each bootstrap sample. This classifier was chosen to address the issue of imbalanced data in the prediction model.

3.4. Model Evaluation. Model performance was evaluated using stratified 5-fold cross-validation with 20 repeats. This approach is appropriate for small-sized datasets and helps eliminate bias and overfitting in model training [66]. In this implementation, the dataset is split into 5 stratified folds with the same ratio of positive and negative cases, and each time one of the folds is used as a test set, and the remaining folds are used for training. This process is repeated 20 times, which provides a more robust model assessment.

To evaluate our models, several performance metrics were used, such as AUC (area under the curve), overall accuracy, recall, and $F2$. AUC is the area under the receiver operating characteristic (ROC) curve that measures the ability of a classifier to correctly distinguish between classes. The $F2$ is a variation of the F -score measure, which calculates the harmonic mean of precision and recall, as shown in equation (1). A high value of the F -score indicates a high balanced classifier performance. In $F2$ calculation, higher weight is given to recall because correctly identifying the minority class (i.e., critical near-crashes) is more important than incorrectly classifying a noncritical conflict as critical. Other metrics such as the overall accuracy alone may not be suitable for model evaluation as they result in poor classification performance for the minority class. Therefore, we opted to use $F2$ as the primary metric for comparing results between different classifiers, though other metrics were also used for comparison:

$$F2 = \frac{(1 + 2^2) \times \text{precision} \times \text{recall}}{2^2 \times \text{precision} + \text{recall}}. \quad (1)$$

4. Results and Discussion

A total of 324 vehicle-bicycle interactions were identified with a T_2 of less than 5 s. Of these events, 85 cases had PET values less than 5 s, and 58 cases had PET values less than 3 s. Considering the T_2 threshold of 3 s reduced the number of events to 174 interactions, of which 56 had a PET value of less than 5 s and 46 had a PET value of less than 3 s. The decrease in the number of cases is due to the fact that a smaller threshold would naturally result in fewer observations.

Figure 5 shows the $F2$ performance of all scenarios of near-crash prediction models. The balanced random forest surpassed SVM and logistic regression in all scenarios. Due to the random undersampling of the majority class on each

bootstrap, the balanced random forest performed better on the imbalanced data. Also, it could be seen that for interactions with T_2 less than 5 s, models using 10-frame monitoring periods had higher performance, while for interactions with T_2 less than 3 s, 20-frame monitoring period models performed better. By increasing the monitoring period length, there are more data available for the model, which potentially could help to predict near-crashes more accurately. However, by increasing the length of the monitoring period, data from way before the near-crash occurrence are added to the model, which could negatively impact the model performance. Therefore, a small trade-off between model performance and monitoring period length is noticeable, showing models with monitoring period lengths of 10 and 20 frames performed slightly better than those with lower or higher monitoring period lengths.

Table 6 shows the detailed result of the best near-crash prediction model for each of the four scenarios.

In order to reduce overfitting and improve model performance, a sequential backward and forward feature selection was performed. These methods were implemented using the Python `Mlxtend` package [67]. In the backward feature selection, the algorithm starts with all of the features, and then in each step, one feature that after elimination maximizes the model performance is removed. Also, in each step, algorithms try adding previously removed features back to the feature subset to see if they could increase the performance. The algorithm went through all 30 features in this study and identified the best subset of features with the highest $F2$ performance in the balanced random forest using the same cross-validation technique. The forward selection utilizes a similar procedure, except that it starts, with zero features, and features are added in each step. In case both forward and backward feature selections reach the same result, the feature subset containing fewer variables is selected for model training to reduce overfitting. Figure 6 shows the best performance obtained from feature selection for each of the scenarios.

Table 7 shows the detailed result of near-crash prediction using the selected subset of features. As can be seen from the table, feature selection significantly improved model performance.

Changing the T_2 threshold from 5 to 3 s resulted in $F2$ improvements, as shown in Table 7. Events with smaller values of T_2 present situations where road users are closer to the TIP. Therefore, the improvements could be attributed to the fact that as road-users move closer to the TIP, their behaviors are better reflected in their kinematic and SSM-related features. Furthermore, lowering the T_2 threshold results in exclusion of less severe interactions and improves the data imbalance, which in turn could improve the model performance.

Lastly, the prediction models were investigated only for the rear-end type of near-crashes. Rear-end near-crashes were identified as interactions with the direction of conflict of one, which almost constitutes half of the interactions in each scenario. Table 8 shows the best results obtained in each scenario for rear-end near-crashes after the feature selection. Comparing Table 8 with Table 7, the near-crash prediction

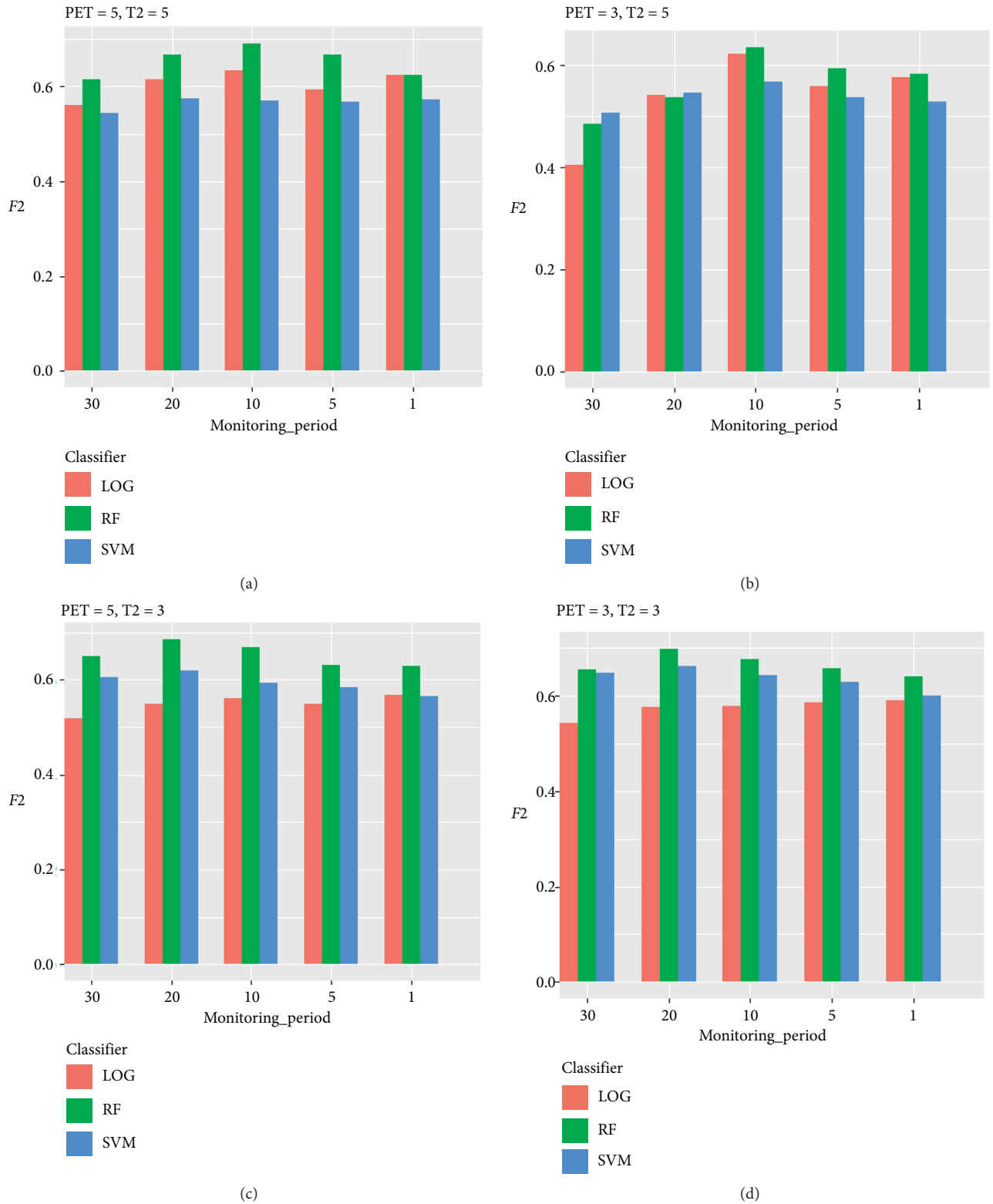


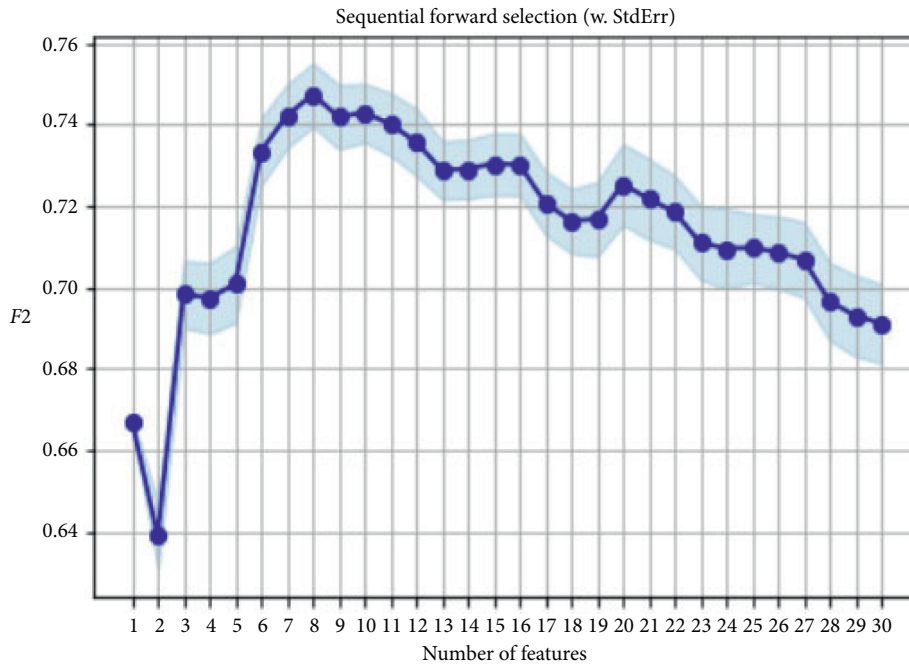
FIGURE 5: F2 performance of all models. (a) PET = 5 and T₂ = 5; (b) PET = 3 and T₂ = 5; (c) PET = 5 and T₂ = 3; (d) PET = 3 and T₂ = 3.

models perform better in almost all of the metrics when considering rear-end interactions alone. The increase in performance could be due to the fact that, in rear-end near-crashes, the interacting road users are on the same or

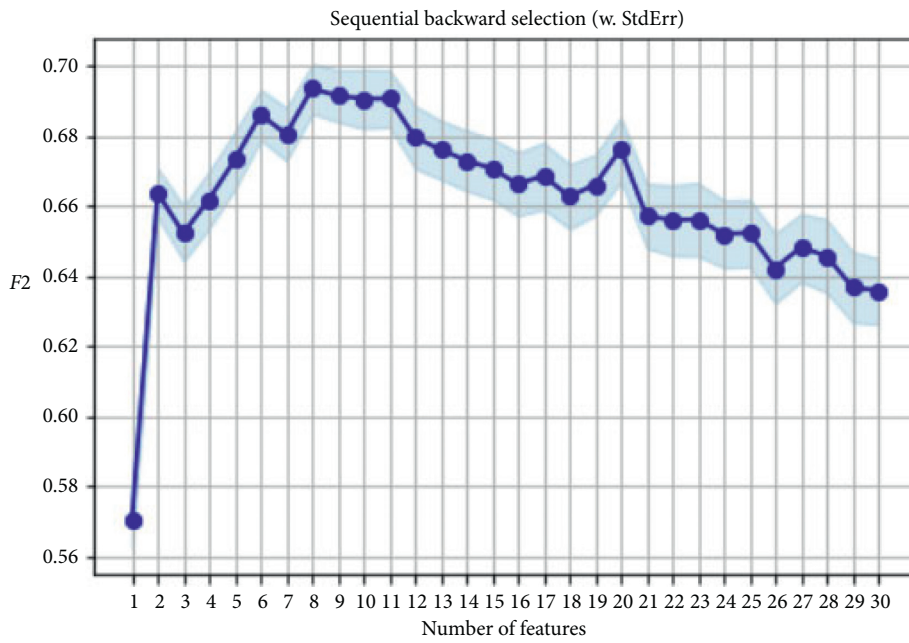
adjacent lanes traveling towards the same direction. Therefore, the future trajectories of these road users are more predictable. In contrast, in other types of near-crashes, such as crossing or sideswipe interactions, it is more challenging

TABLE 6: Result of the best near-crash prediction models using all features.

Scenarios	Monitoring period	Classifier	Accuracy	AUC	Recall	F2	Precision
PET = 5 and T ₂ = 5	10	RF	0.698	0.784	0.778	0.691	0.488
PET = 3 and T ₂ = 5	10	RF	0.657	0.776	0.824	0.636	0.338
PET = 5 and T ₂ = 3	20	RF	0.684	0.768	0.745	0.686	0.537
PET = 3 and T ₂ = 3	20	RF	0.675	0.785	0.806	0.699	0.478

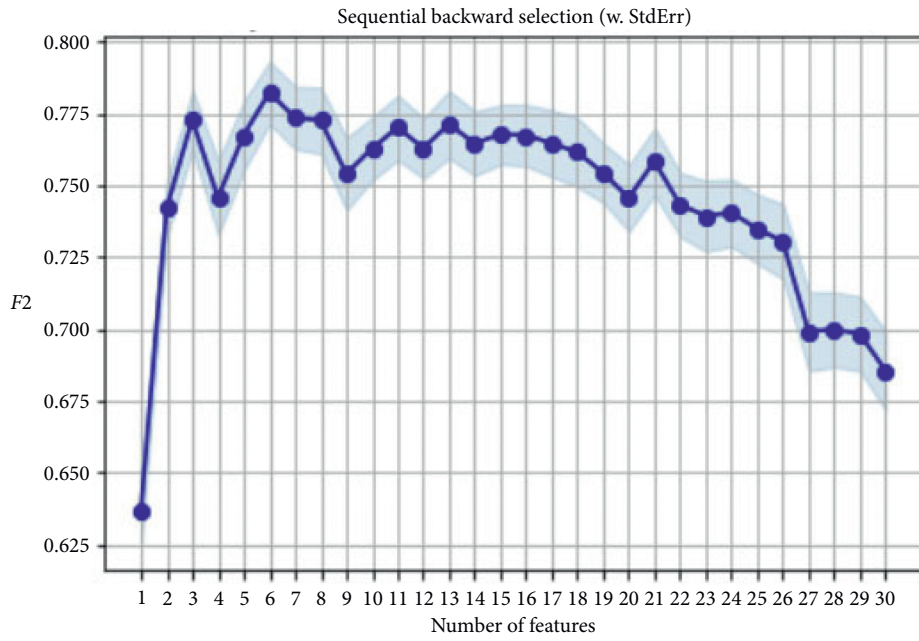


(a)

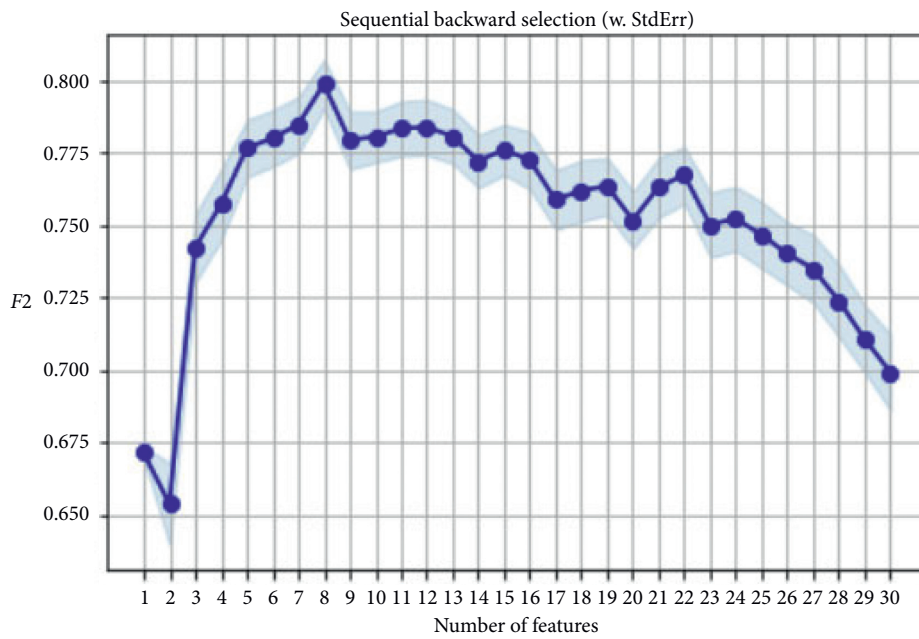


(b)

FIGURE 6: Continued.



(c)



(d)

FIGURE 6: Best result achieved using the forward and backward feature selection in each scenario. (a) Monitoring period = 10, PET = 5, and $T_2 = 5$; (b) monitoring period = 10, PET = 3, and $T_2 = 5$; (c) monitoring period = 20, PET = 5, and $T_2 = 3$; (d) monitoring period = 20, PET = 3, and $T_2 = 3$.

TABLE 7: Result of the best near-crash prediction models after feature selection.

Scenarios	Monitoring period	Classifier	Accuracy	AUC	Recall	F2	Precision
PET = 5 and $T_2 = 5$	10	RF	0.708	0.807	0.87	0.755	0.501
PET = 3 and $T_2 = 5$	10	RF	0.651	0.802	0.906	0.681	0.346
PET = 5 and $T_2 = 3$	20	RF	0.744	0.829	0.845	0.774	0.597
PET = 3 and $T_2 = 3$	20	RF	0.714	0.814	0.91	0.783	0.514

TABLE 8: Result of the best near-crash prediction models for rear-end interactions after feature selection.

Scenarios	Monitoring period	Classifier	Accuracy	AUC	Recall	<i>F2</i>	Precision
PET = 5 and $T_2 = 5$	10	RF	0.766	0.802	0.936	0.841	0.614
PET = 3 and $T_2 = 5$	10	RF	0.767	0.873	0.92	0.787	0.516
PET = 5 and $T_2 = 3$	20	RF	0.802	0.865	0.953	0.87	0.671
PET = 3 and $T_2 = 3$	20	RF	0.844	0.922	0.941	0.878	0.734

to predict the road-user movements. The sudden changes in the direction of travel and yielding for traffic when making turns or crossing add more uncertainty to the prediction problem.

5. Conclusions and Future Works

There has been an alarming rise in the number of crashes involving bicyclists at signalized intersections. This study adopted a proactive safety assessment approach to predict critical bicycle-vehicle near-crashes at signalized intersections a few seconds before they occur. Kinematic data of interacting road users and SSMs during a monitoring period were used to develop prediction models to distinguish critical from noncritical events. Video data collected from ten signalized intersections in the city of San Diego were used in this study. Critical interactions between vehicles and bicyclists were labeled using two PET thresholds of 3 and 5 s. Different monitoring periods were also defined using two T_2 thresholds of 3 and 5 s as well as five different monitoring period lengths from 1/30 to 1 second. Near-crash prediction models were developed and compared using three machine learning classifiers of SVM, logistic regression, and random forest.

In almost all metrics, the balanced random forest provided higher prediction performance compared to SVM and logistic regression and, therefore, was identified as the best classifier. While all four scenarios of near-crash prediction models performed fairly well with recalls above 70% and *F2* above 60%, two feature selection methods of backward and forward were adopted to reduce the overfitting and improve the model performance. These two algorithms were implemented in order to find the subset of features with the highest *F2* and with the same cross-validation as for model training. The feature selection significantly improved the recall and *F2* metrics to above 85% and around 70% across all four scenarios, respectively. Also, the AUC metric for all the models was above 80% after feature selection. Predicting rear-end near-crashes recall values were further improved to above 90% in all the scenarios, suggesting better prediction performance can be achieved when only a certain conflict type is taken into account. The results also showed that when the end of the monitoring period is defined with a T_2 threshold of 5 s, 10-frame monitoring periods led to the best prediction performance. When a T_2 threshold of 3 s was used, 20-frame monitoring periods resulted in the best performance. Monitoring periods smaller or larger than 10 and 20 led to lower prediction performance in all scenarios. This suggested a small trade-off between the monitoring period length and the prediction performance.

The models developed in this study could be utilized in intelligent transportation management centers focusing on proactive safety applications. The modeling framework used in this study could also be extended and applied to conflicts involving other road-user types (e.g., vehicle-scooter). In a connected environment, these models can inform proactive safety planning in real time by identifying where and when critical conflicts occur between road users. Developing accurate near-crash prediction models for all types of interactions is challenging. Typically, cyclists move faster than pedestrians, and their movements are generally more unpredictable than vehicles. Moreover, many bicyclists tend to participate in dangerous and sudden maneuvers while passing the intersection. The emerging transportation technology such as automated vehicles, as well as newer modes of transportation such as electric bicycles and scooters with unknown behavioral impacts, would also add to the complexity of the prediction problem. Future research may focus on extensive data collection for all road-user types and at different geographical locations in order to obtain representative interaction samples. The present study focused on developing models using kinematic data only. However, it should be noted that other factors such as geometric design, traffic signal schemes, and weather could also play important roles in identifying the severity of near-crash events. While this study adopted SSM thresholds recommended by previous studies, it should be noted that any SSM and threshold should be validated as there is no consensus among researchers on what metrics or thresholds are the right ones for unsafe event identification. Future research may focus on including other environmental and design characteristics in near-crash identification and also incorporating more metrics and thresholds as well as qualitative analysis to conduct validation and sensitivity analysis studies. Future work could also entail investigating multiple conflict severity classes, other methods, and features.

Data Availability

The data used to support the findings of this study are part of a research project and cannot be shared due to privacy reasons.

Disclosure

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

Funding for this research was partly provided by the State of California's Senate Bill 1 "The Road Repair and Accountability Act of 2017" and administered by the California State University Transportation Consortium. This research was also partly funded by the Safety through Disruption (Safe-D) National University Transportation Center (UTC), a grant from the U.S. Department of Transportation's University Transportation Centers Program (Federal Grant number: 69A3551747115).

References

- [1] U.S. Department of Transportation, *National Highway Traffic Safety Administration: 2018 Fatal Motor Vehicle Crashes: Overview*, U.S. Department of Transportation, Washington, DC, USA, 2020.
- [2] N. Bhuiyan, E. Parentela, and V. S. Inapuri, "Analysis of signalized intersection crashes," in *Institute of Transportation Engineers—Western District Meeting of the Institute of Transportation Engineers*, Albuquerque, New Mexico, July 2016.
- [3] A. A. Workineh, *Analysis of the Relationship between Traffic Conflicts and Level of Service at Four-Legged, Signalized Intersections in Sacramento*, Thesis, California State University, Sacramento, CA, USA, 2014.
- [4] Å. Svensson and C. Hydén, "Estimating the severity of safety related behaviour," *Accident Analysis & Prevention*, vol. 38, no. 2, pp. 379–385, 2006.
- [5] M. Hasani, A. Jahangiri, I. N. Sener et al., "Identifying high-risk intersections for walking and bicycling using multiple data sources in the city of San Diego," *Journal of Advanced Transportation*, vol. 2019, Article ID 9072358, 2019.
- [6] P. Chen, W. Zeng, G. Yu, and Y. Wang, "Surrogate safety analysis of pedestrian-vehicle conflict at intersections using unmanned aerial vehicle videos," *Journal of Advanced Transportation*, vol. 2017, Article ID 5202150, 2017.
- [7] A. Laureshyn, Å. Svensson, and C. Hydén, "Evaluation of traffic safety, based on micro-level behavioural data: theoretical framework and first implementation," *Accident Analysis & Prevention*, vol. 42, no. 6, pp. 1637–1646, 2010.
- [8] D. Lord and B. N. Persaud, "Estimating the safety performance of urban road transportation networks," *Accident Analysis & Prevention*, vol. 36, no. 4, pp. 609–620, 2004.
- [9] S. G. Machiani and M. Abbas, "Safety surrogate histograms (SSH): a novel real-time safety assessment of dilemma zone related conflicts at signalized intersections," *Accident Analysis & Prevention*, vol. 96, pp. 361–370, 2016.
- [10] N. K. Salman and K. J. Al-Maita, "Safety evaluation at three-leg, unsignalized intersections by traffic conflict technique," no. 1485, , pp. 177–185, *Transportation Research Record*, Thousand Oaks, CA, USA, 1995.
- [11] T. Sayed and S. Zein, "Traffic conflict standards for intersections," *Transportation Planning and Technology*, vol. 22, no. 4, pp. 309–323, 1999.
- [12] S. R. Perkins and J. I. Harris, *Traffic Conflict Characteristics: Freeway Curve and Exit Area FI*, December, 1966, General Motors Corporation, Detroit, Michigan, USA, 1967.
- [13] T. W. Forbes, *Analysis of "Near Accident" Reports*, vol. 152, pp. 23–37, *Highway Research Board Bulletin*, Washington, DC, USA, 1957.
- [14] D. Gettman and L. Head, "Surrogate safety measures from traffic simulation models," *Transportation Research Record*, vol. 1840, no. 1, pp. 104–115, 2003.
- [15] A. Tarko, G. Davis, N. Saunier, T. Sayed, and S. Washington, *Surrogate Measures of Safety-White paper*, Transportation Research Board, Washington, DC, USA, 2009.
- [16] F. Amundsen and C. Hyden, *Proceedings of First Workshop on Traffic Conflicts*. Oslo, TTI, Oslo, Norway and LTH Lund University, Lund, Sweden, 1977.
- [17] C. Hydén, *The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique*, Bulletin Lund Institute of Technology, Lund, Sweden, 1987.
- [18] A. Varhelyi, A. Laureshyn, C. Johnsson et al., *Surrogate safety measures and traffic conflict observations, in How to Analyse Accident Causation?: A Handbook with Focus on Vulnerable Road Users*, E. Polders and T. Brijs, Eds., pp. 95–128, InDeV, Horizon 2020 project, Hasselt University, Brussels, Belgium, 1st edition, 2018, [https://portal.research.lu.se/portal/en/publications/surrogate-safety-measures-and-traffic-conflict-observations\(ef283c81-c2f8-4a81-8f4b-06a512ba3687\)/export.html](https://portal.research.lu.se/portal/en/publications/surrogate-safety-measures-and-traffic-conflict-observations(ef283c81-c2f8-4a81-8f4b-06a512ba3687)/export.html).
- [19] M. H. Zaki, T. Sayed, A. Tageldin, and M. Hussein, "Application of computer vision to diagnosis of pedestrian safety issues," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2393, no. 1, pp. 75–84, 2013.
- [20] J. C. Hayward, "Near miss determination through use of a scale of danger," in *51st Annual Meeting of the Highway Research Board*, Washington, DC, USA, 1972.
- [21] B. L. Allen, B. T. Shin, and P. J. Cooper, *Analysis of Traffic Conflicts and Collisions*, Transportation Research Record, Thousand Oaks, CA, USA, 1978.
- [22] M. Hossain and Y. Muromachi, "A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways," *Accident Analysis & Prevention*, vol. 45, pp. 373–381, 2012.
- [23] A. Abdulhafedh, "Crash frequency analysis," *Journal of Transportation Technologies*, vol. 6, no. 4, pp. 169–180, 2016.
- [24] F. L. Mannering and C. R. Bhat, "Analytic methods in accident research: methodological frontier and future directions," *Analytic Methods in Accident Research*, vol. 1, pp. 1–22, 2014.
- [25] J. C. Milton, V. N. Shankar, and F. L. Mannering, "Highway accident severities and the mixed logit model: an exploratory empirical analysis," *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 260–266, 2008.
- [26] C. Dong, D. B. Clarke, X. Yan, A. Khattak, and B. Huang, "Multivariate random-parameters zero-inflated negative binomial regression model: an application to estimate crash frequencies at intersections," *Accident Analysis & Prevention*, vol. 70, pp. 320–329, 2014.
- [27] L.-Y. Chang, "Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network," *Safety Science*, vol. 43, no. 8, pp. 541–557, 2005.
- [28] X. Li, D. Lord, Y. Zhang, and Y. Xie, "Predicting motor vehicle crashes using support vector machine models," *Accident Analysis & Prevention*, vol. 40, no. 4, pp. 1611–1618, 2008.
- [29] H. Huang, Q. Zeng, X. Pei, S. C. Wong, and P. Xu, "Predicting crash frequency using an optimised radial basis function neural network model," *Transportmetrica A: Transport Science*, vol. 12, no. 4, pp. 330–345, 2016.
- [30] J.-S. Oh, C. Oh, S. G. Ritchie, and M. Chang, "Real-time estimation of accident likelihood for safety enhancement," *Journal of Transportation Engineering*, vol. 131, no. 5, pp. 358–363, 2005.

- [31] M. Hossain and Y. Muromachi, "Development of a real-time crash prediction model for urban expressway," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 8, pp. 2092–2107, 2010.
- [32] J. Sun and J. Sun, "Real-time crash prediction on urban expressways: identification of key variables and a hybrid support vector machine model," *IET Intelligent Transport Systems*, vol. 10, no. 5, pp. 331–337, 2016.
- [33] J. Yuan and M. Abdel-Aty, "Approach-level real-time crash risk analysis for signalized intersections," *Accident Analysis & Prevention*, vol. 119, pp. 274–289, 2018.
- [34] K. El-Basyouny and T. Sayed, "Safety performance functions using traffic conflicts," *Safety Science*, vol. 51, no. 1, pp. 160–164, 2013.
- [35] E. Sacchi and T. Sayed, "Conflict-based safety performance functions for predicting traffic collisions by type," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2583, no. 1, pp. 50–55, 2016.
- [36] T. Fu, W. Hu, L. Miranda-Moreno, and N. Saunier, "Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 222–240, 2019.
- [37] N. Saunier and T. Sayed, "Probabilistic framework for automated analysis of exposure to road collisions," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2083, no. 1, pp. 96–104, 2008.
- [38] N. Saunier, T. Sayed, and K. Ismail, "Large-scale automated analysis of vehicle interactions and collisions," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2147, no. 1, pp. 42–50, 2010.
- [39] M. R. Parker and C. V. Zegeer, *Traffic Conflict Techniques for Safety and Operations: Engineers Guide*, Final Report (No. FHWA/IP-88/026), 1989, <https://trid.trb.org/view/1183295>.
- [40] Z. He, X. Qin, P. Liu, and M. A. Sayed, "Assessing surrogate safety measures using a safety pilot model deployment dataset," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 38, pp. 1–11, 2018.
- [41] L. Zheng, K. Ismail, and X. Meng, "Freeway safety estimation using extreme value theory approaches: a comparative study," *Accident Analysis & Prevention*, vol. 62, pp. 32–41, 2014.
- [42] L. Zheng, K. Ismail, T. Sayed, and T. Fatema, "Bivariate extreme value modeling for road safety estimation," *Accident Analysis & Prevention*, vol. 120, pp. 83–91, 2018.
- [43] H. Farah and C. L. Azevedo, "Safety analysis of passing maneuvers using extreme value theory," *IATSS Research*, vol. 41, no. 1, pp. 12–21, 2017.
- [44] C. Wang, C. Xu, and Y. Dai, "A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data," *Accident Analysis & Prevention*, vol. 123, pp. 365–373, 2019.
- [45] L. Zheng, T. Sayed, M. Essa, and Y. Guo, "Do simulated traffic conflicts predict crashes? an investigation using the extreme value approach," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 631–636, Auckland, New Zealand, October 2019.
- [46] L. Zheng and T. Sayed, "Comparison of traffic conflict indicators for crash estimation using peak over threshold approach," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 5, pp. 493–502, 2019.
- [47] F. Guo, S. G. Klauer, J. M. Hankey, and T. A. Dingus, "Near crashes as crash surrogate for naturalistic driving studies," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2147, no. 1, pp. 66–74, 2010.
- [48] B. Cheng, Q. Lin, T. Song, Y. Cui, L. Wang, and S. Kuzumaki, "Analysis of driver brake operation in near-crash situation using naturalistic driving data," *International Journal of Automotive Engineering*, vol. 2, no. 4, pp. 87–94, 2011.
- [49] K.-F. Wu, J. Aguero-Valverde, and P. P. Jovanis, "Using naturalistic driving data to explore the association between traffic safety-related events and crash risk at driver level," *Accident Analysis & Prevention*, vol. 72, pp. 210–218, 2014.
- [50] N. Formosa, M. Quddus, S. Ison, M. Abdel-Aty, and J. Yuan, "Predicting real-time traffic conflicts using deep learning," *Accident Analysis & Prevention*, vol. 136, 2020.
- [51] O. A. Osman, M. Hajji, P. R. Bakhit, and S. Ishak, "Prediction of near-crashes from observed vehicle kinematics using machine learning," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 12, 2019.
- [52] M. Essa and T. Sayed, "Full Bayesian conflict-based models for real time safety evaluation of signalized intersections," *Accident Analysis & Prevention*, vol. 129, pp. 367–381, 2019.
- [53] Y. Ma, H. Meng, S. Chen, J. Zhao, S. Li, and Q. Xiang, "Predicting traffic conflicts for expressway diverging areas using vehicle trajectory data," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 3, 2020.
- [54] M. Essa and T. Sayed, "Traffic conflict models to evaluate the safety of signalized intersections at the cycle level," *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 289–302, 2018.
- [55] L. N. Peesapati, M. P. Hunter, and M. O. Rodgers, "Evaluation of postencroachment time as surrogate for opposing left-turn crashes," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2386, no. 1, pp. 42–51, 2013.
- [56] A. Svensson, *A Method for Analysing the Traffic Process in a Safety Perspective*, Lund Institute of Technology, Sweden, 1998.
- [57] S. M. S. Mahmud, L. Ferreira, M. S. Hoque, and A. Tavassoli, "Application of proximal surrogate indicators for safety evaluation: a review of recent developments and research needs," *IATSS Research*, vol. 41, no. 4, pp. 153–163, 2017.
- [58] K. Xie, D. Yang, K. Ozbay, and H. Yang, "Use of real-world connected vehicle data in identifying high-risk locations based on a new surrogate safety measure," *Accident Analysis & Prevention*, vol. 125, pp. 311–319, 2019.
- [59] T. Sayed, M. H. Zaki, and J. Autey, "Automated safety diagnosis of vehicle-bicycle interactions using computer vision analysis," *Safety Science*, vol. 59, pp. 163–172, 2013.
- [60] G. Grayson, C. K. Hyden, and J. H. Kraay, N. Muhrad and S. Oppe, *The Malmö Study: A Calibration of Traffic Conflict Techniques. A Study organised by ICTCT (The International Committee on Traffic Conflict Techniques)*, Institute for Road Safety Research SWOV, The Netherlands, 1984.
- [61] S. Zangenehpour, L. F. Miranda-Moreno, and N. Saunier, "Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: methodology and application," *Transportation Research Part C: Emerging Technologies*, vol. 56, pp. 161–176, 2015.
- [62] S. Zangenehpour, J. Strauss, L. F. Miranda-Moreno, and N. Saunier, "Are signalized intersections with cycle tracks safer? a case-control study based on automated surrogate safety analysis using video data," *Accident Analysis & Prevention*, vol. 86, pp. 161–172, 2016.

- [63] S. Kothuri, E. Smaglik, A. Kading et al., *Addressing Bicycle-Vehicle Conflicts with Alternate Signal Control Strategies*, Portland State University, Portland, Oregon, 2018.
- [64] J. Stipancic, S. Zangenehpour, L. Miranda-Moreno, N. Saunier, and M.-A. Granié, "Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology," *Accident Analysis & Prevention*, vol. 97, pp. 19–27, 2016.
- [65] A. Jahangiri, A. Katthe, A. Sohrabi et al., *Developing a Computer Vision-Based Decision Support System for Intersection Safety Monitoring and Assessment of Vulnerable Road Users*, Mineta Transportation Institute Publications, San Jose, CA, United States, 2020.
- [66] G. C. Cawley and N. L. Talbot, "On over-fitting in model selection and subsequent selection bias in performance evaluation," *The Journal of Machine Learning Research*, vol. 11, pp. 2079–2107, 2010.
- [67] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, vol. 15, no. 11, pp. 1119–1125, 1994.

Research Article

Analysis of Electric Vehicle Charging Behavior Patterns with Function Principal Component Analysis Approach

Chenxi Chen,¹ Yang Song,¹ Xianbiao Hu ,¹ and Ivan G. Guardiola ²

¹Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA

²School of Business Administration, Fort Lewis College, 1000 Rim Drive, Durango, CO 81301, USA

Correspondence should be addressed to Xianbiao Hu; xbhu@mst.edu

Received 8 May 2020; Revised 9 November 2020; Accepted 11 November 2020; Published 23 November 2020

Academic Editor: Yuanyuan Zhang

Copyright © 2020 Chenxi Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This manuscript focused on analyzing electric vehicles' (EV) charging behavior patterns with a functional data analysis (FDA) approach, with the goal of providing theoretical support to the EV infrastructure planning and regulation, as well as the power grid load management. 5-year real-world charging log data from a total of 455 charging stations in Kansas City, Missouri, was used. The focuses were placed on analyzing the daily usage occupancy variability, daily energy consumption variability, and station-level usage variability. Compared with the traditional discrete-based analysis models, the proposed FDA modeling approach had unique advantages in preserving the smooth function behavior of the data, bringing more flexibility in the modeling process with little required assumptions or background knowledge on independent variables, as well as the capability of handling time series data with different lengths or sizes. In addition to the patterns revealed in the EV charging station's occupancy and energy consumption, the differences between EV driver's charging time and parking time were analyzed and called for the needs for parking regulation and enforcement. The different usage patterns observed at charging stations located on different land-use types were also analyzed.

1. Introduction

Electric vehicles (EVs) produce fewer emissions that contribute to climate change and smog than conventional vehicles and help the United States achieve a greater diversity of fuel choices available for transportation. The evolution of EVs has advanced from models best suited for commuting or traveling short distances to vehicles that can travel more than 200 or even 300 miles per charge.

Proper planning of the EV charging infrastructure and scientific determination of their locations are critical to promoting EV ownership and usage. Modeling efforts can be found in the literature, such as the electric vehicle infrastructure projection (EVI-Pro) model developed by the National Renewable Energy Lab to address the fundamental question of how much charging infrastructure is needed in the United States to support Plug-in-EVs (PEVs) [1]. The model generated a quantitative estimate for a US network of

nonresidential (public and workplace) EVSE that would be needed to support broader PEV adoption. He et al. studied how to optimally locate public charging stations on a road network, considering drivers' spontaneous adjustments and interactions of travel and recharging decisions [2]. A bilevel programming model with the consideration of EV's driving range was proposed in [3], with the upper level to optimize the position of charging stations so as to maximize the path flows that used the charging stations, while the user equilibrium of route choice with the EV's driving range constraint was formulated in the lower level. Other research on EV charging station locations can also be found in [4, 5] and many others.

Another approach to supporting the planning of charging infrastructure was to perform analysis of EV-related data, with the goal of identifying charging behavior patterns and inferring the scenarios of when and where people need to charge their vehicles. For example, the

driving data in Denmark was analyzed to extract the information of driving distances and driving time periods which were used to represent the driving requirements and the EV unavailability. The Danish National Transport Survey data were used to implement the driving data analysis [6]. The analysis of charge event data in Ireland for public charging infrastructure, including data from fast-charging infrastructure, and additionally a limited quantity of household data was performed in [7]. Sun et al. studied driver's charging timing decisions, in which a mixed logit model with unobserved heterogeneity is applied to panel data extracted from a two-year field trial on battery electric vehicle usage in Japan [8]. The analysis over the real-world dataset can also be found in [9, 10] and others.

This manuscript focused on performing analysis over the 5-year real-world charging event log data, from a total of 455 charging stations in Kansas City, Missouri (KCMO), with a functional data analysis (FDA) approach. The EV charging equipment recorded which vehicle was charged at which charging station, at what day and time. Such charging event log data contained many significant pieces of information for understanding EV charging patterns and user behavior. The goal of this research was to provide theoretical support to the EV infrastructure planning and regulation, as well as the power grid load management. We argue that compared with the existing research over the real-world charging event data, the proposed FDA modeling approach had many unique advantages over the prevailing discrete-based analysis models and led to some important insights that were difficult to model or discover with the other approaches.

Commonly, time series data (such as the EV charging log data used in this research) were treated as multivariate data because they were given as a finite discrete time series [11–13]. This usual multivariate approach completely ignored important information about the smooth functional behavior of the generating process that underpins the data [14]. For example, in our context, the vehicles' charging process was continuous and so was the time-dependent occupancy of a particular charging station. Additionally, in the previous research, performance measurements need to be defined by the researchers to extract useful information from the raw dataset, before any meaningful analysis can be performed. However, they were usually defined arbitrarily, based on the researcher's experience in the field. Instead of assuming a variety of explanatory variables, which was difficult or even impossible to enumerate and collect data for, FDA is much more flexible with little required assumptions or background knowledge on independent variables. Last not but least, time series data often has different time intervals or different lengths which are hard to deal with by other tools. In our context, some charging stations were more frequently used and might have thousands of charging records while others might only have a few hundred. It was thus impossible to apply principal component analysis (PCA) to the charging log dataset directly because of the dimension inconsistency.

The basic idea behind FDA is to express discrete observations arising from time series in the form of a function (i.e., to create functional data) that represents the entire

measured function as a single observation and then to draw the modeling and/or prediction information from a collection of functional data by applying statistical concepts from multivariate data analysis [15]. With this said, this manuscript firstly represented the EV charging dataset with a continuous functional form, then performed function principal component (FPC) analysis to identify the main contributing principal components (PC), and analyzed the dataset from different perspectives to understand EV owner's charging behavior patterns.

This research aimed to provide theoretical support to the EV infrastructure planning and regulation, as well as the power grid load management. To achieve such goals, the focuses were placed on three aspects. (1) The first aspect is the variability analysis of the daily usage patterns of all EV charging stations, in which the 24-hour occupancy of all charging stations in one day was treated as one continuous curve. Such analysis can provide insights and directly support the planning of new EV charging infrastructures. (2) The second aspect is the variability analysis of the daily energy consumption of all EV charging stations, in which the total energy consumption of all charging stations in one day was treated as one continuous curve. Such analysis was important from the power grid load management perspective. (3) At the station level, the usage pattern variabilities were analyzed, in which one station's usage over the entire observation period was treated as a continuous curve. This analysis revealed insights on the usage pattern differences at the station level and was combined with the land-use information for better EV charging infrastructure planning and management purposes.

The remaining part of this paper is organized as follows. The charging event log dataset used in this research is firstly presented in Section 2. Section 3 presents the analysis methodology, including the data smoothing, variable calculation, and the functional principal component analysis. The analysis results are shown and compared in Section 4. Section 5 concludes this research.

2. Data

This section presents the real-world charging event log data used in this research. The data was collected from 455 charging stations between January 2014 and November 2019 in Kansas City, Missouri (KCMO). The dataset included a total of 226,652 charging records from 4,921 users. Most of the stations were concentrated in the downtown area of KCMO. The spatial distribution of charging stations was shown in Figure 1, in which Figure 1(a) showed an overview and Figure 1(b) zoomed in to the downtown area.

In the collected dataset, each row contained the information of a charging event and had a total of 30 columns/attributes. Table 1 showed the sample data from the dataset, in which only the most critical and relevant information was displayed. The complete dataset included information of the following three categories:

- (1) Charging station information: including a unique station ID, station name, address and zip code where

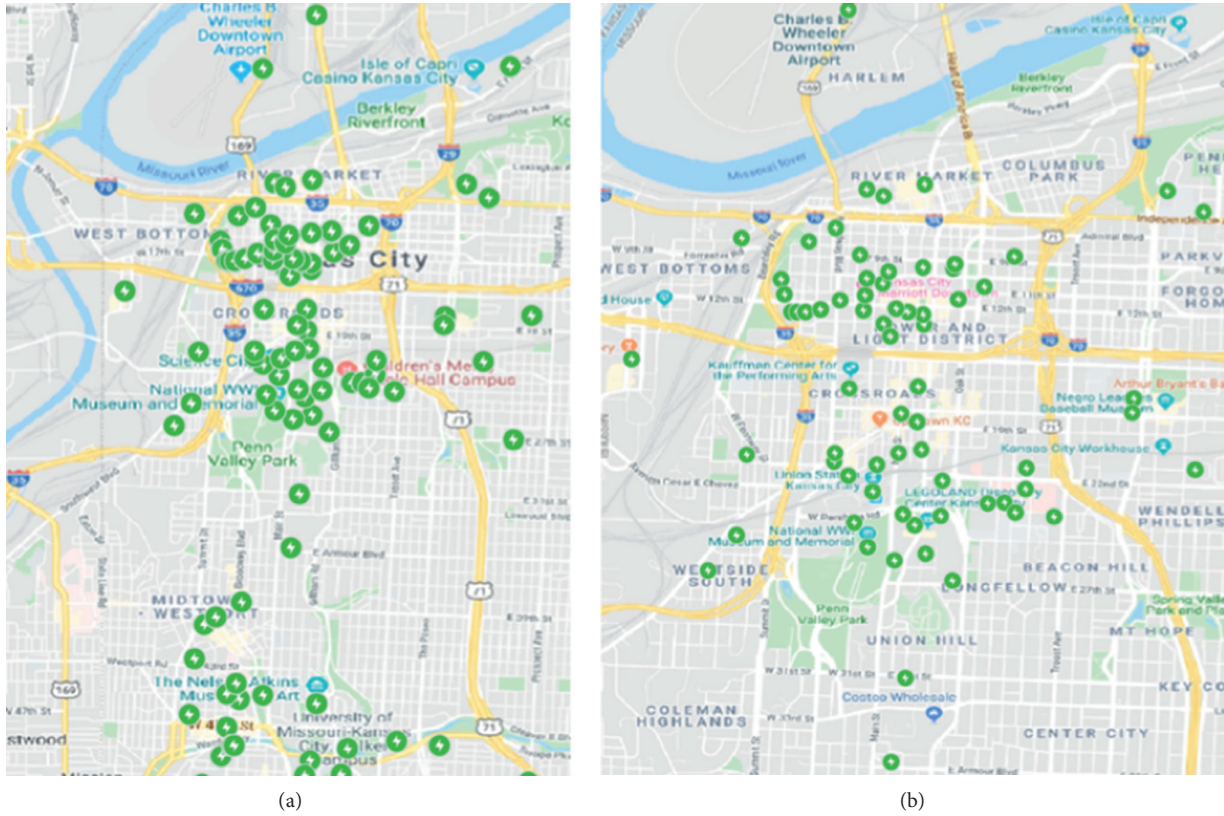


FIGURE 1: Locations of the charging station in KCMO. (a) Overview. (b) Downtown.

TABLE 1: Sample Charging Event Log data.

Station name	Start date	End date	Time duration (hh: mm:ss)	Charging time (hh: mm:ss)	Energy (kWh)	Latitude	Longitude	User ID
KCPL/@JE DUNN	9/18/2018	9/18/2018	1:23:15	1:23:02	5.934	39.1011	94.5763	756927
PG125 C	14:27	15:50						
KCPL/@WOLF PG-129A	9/18/2018	9/18/2018	8:42:04	2:15:18	6.766	39.0999	94.5791	1533491
KCPL/@LOOSE	9/18/2018	9/18/2018	0:00:30	0:00:00	0	39.0346	94.5947	598567
PRK-121A	15:09	15:10						

the station was located at, MAC address, latitude and longitude of the station, and type of the charging ports which included level 1, level 2, and DC fast charge

- (2) Electric vehicle attribute: including a unique ID of the electric vehicle and zip code where this electric was registered in (which is usually the zip code of the driver’s home)
- (3) Charging event data: including the start date and time of the charging event, end date and time of the charging event, charging time which is equal to the end time minus start time, total duration which included not only the time spent on charging but also the time spent on parking afterward, start state of charge (SOC), end SOC, energy charged, Greenhouse Gas (GHG) saving, and information on how was the charging event ended (e.g., terminated by

customer or server). Duration is the total time that a station is occupied, which is one of the most significant properties we are interested

3. Methodology

This section presents the analysis methodology used in this manuscript, including a brief overview of the function data analysis approach, charging pattern definition, and functional principal component analysis.

3.1. FDA Method Overview. To process “curve-like” data that are continuous in nature, such as the time-dependent charging station usage rates of this manuscript, one advanced and popular method is functional data analysis [16]. Apart from the commonly seen multivariate data analysis approaches, the proposed FDA approach considered EV

charging usage as a function of time; thus, all the EV charging events that were sampled in different scales, from different charging stations built at different time periods, and used with different frequencies with different data sizes, were all modeled uniformly by functions. In other words, under the functional data analysis approach, each charging pattern to be defined in Section 3.2 was treated as one functional data. By applying basis expansion techniques such as B-spline expansion denoted in (4) [17, 18], each charging pattern can be modeled and expressed in a functional form.

$$y(t) = \sum_{j=1}^M \beta_j \varphi_j(t), \quad (1)$$

where $y(t)$ is the original function, $\varphi_j(t)$ is j^{th} basis functions, and β_j is the coefficient of the corresponding basis function.

With such a data analysis approach, all charging patterns which were sampled in different scales can be uniformly expressed in the same functional form. Additional benefits of such an approach also included the reduction of unnecessary noise in raw data by basis expansion smoothing. Based on this model, all the information from the raw data can be projected to M basis coefficients β_j . Obtaining the basis coefficients can be done through an ordinary least square (OLS) regression. This process is also known as B-spline smoothing. Section 3.3 will further describe the basis expansion and modeling process.

3.2. Charging Pattern Definition. This subsection defines the three charging patterns to be analyzed, corresponding to the three analyses performed in the numeric analysis section.

3.2.1. Daily Usage Occupancy. Daily usage occupancy was defined to measure the 24-hour time-dependent usage occupancy within a single day, by aggregating the charging events at all charging stations.

A binary variable $u_{j,d}(t)$ was firstly defined to denote the usage condition of a charging station j in hour t at day d . If the charging station was used for at least once, $u_{j,d}(t) = 1$, else 0.

$$u_{j,d}(t) = \begin{cases} 1, & \text{the station was in use,} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Next, the 24-hour time-dependent occupancy for each day can be calculated by aggregating all charging station's usage, so that in the end, one curve was generated to represent the daily usage occupancy of each day.

$$u_d(t) = \frac{1}{J} \sum_j u_{j,d}(t), \quad (3)$$

where $u_d(t)$ means the average occupancy of time t at day d , and J means the total number of stations on day d . Note J is day-dependent, so $J = J(d)$.

3.2.2. Daily Energy Consumption. As shown in Table 1, the energy consumption $e_{i,d}$ associated with each charging event i at day d was recorded and thus was directly available. First, $e_{i,d}$ was proportionally assigned to each hour, so that

$$e'_{i,d}(t) = \frac{e_{i,d}}{l_{i,d}} * l_{i,d}(t), \quad (4)$$

in which $e'_{i,d}(t)$ was the proportion of energy consumption $e_{i,d}$ in hour t , $l_{i,d}$ was the duration of charging event i , and $l_{i,d}(t)$ was the proportion of $l_{i,d}$ in hour t .

Next, the 24-hour time-dependent energy consumption for each day can be calculated by doing aggregation over all charging stations, so that in the end, one curve was generated to represent the daily energy consumption of each day.

$$e_d(t) = \frac{1}{J} \sum_i e'_{i,d}(t). \quad (5)$$

3.2.3. Station-Level Occupancy. Similar to the daily usage occupancy calculation, to analyze the difference between stations, aggregations can be performed over the days. For each station j , its aggregated occupancy at time t was denoted as $u_j(t)$ and calculated as follows.

$$u_j(t) = \frac{1}{T} \sum_d u_{j,d}(t), \quad (6)$$

where $u_{j,d}(t)$ was calculated from (2), and T denoted the total number of days in the analyzed time period. In the end, one curve was generated to represent the aggregated usage occupancy of each charging station.

3.3. Data Smoothing. This subsection focuses on how to represent the charging patterns defined above as curves. Since $u_d(t)$, $e_d(t)$, and $u_j(t)$ are all time-dependent, they can be represented by $(t, u_d(t), e_d(t), \text{ and } u_j(t))$. Based on B-spline expansion, these discrete points can be modeled by a continuous function:

$$\begin{aligned} u_d(t) &= \sum_{j=1}^M \alpha_{i,j} \varphi_j(t) = [\alpha_{i,1} \ \alpha_{i,2} \ \dots \ \alpha_{i,M}] \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \dots \\ \varphi_M(t) \end{bmatrix} = \alpha_i^T \phi(t), \\ e_d(t) &= \sum_{j=1}^M \beta_{i,j} \varphi_j(t) = [\beta_{i,1} \ \beta_{i,2} \ \dots \ \beta_{i,M}] \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \dots \\ \varphi_M(t) \end{bmatrix} = \beta_i^T \phi(t), \\ u_j(t) &= \sum_{j=1}^M \gamma_{i,j} \varphi_j(t) = [\gamma_{i,1} \ \gamma_{i,2} \ \dots \ \gamma_{i,M}] \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \dots \\ \varphi_M(t) \end{bmatrix} = \gamma_i^T \phi(t). \end{aligned} \quad (7)$$

An example of B-spline expansion was depicted in Figure 2, where a smoothed function (solid black curve) was represented as a summation of B-spline basis functions (dashed black curves) to model the raw daily usage occupancy data (red diamond). The heights of these basis

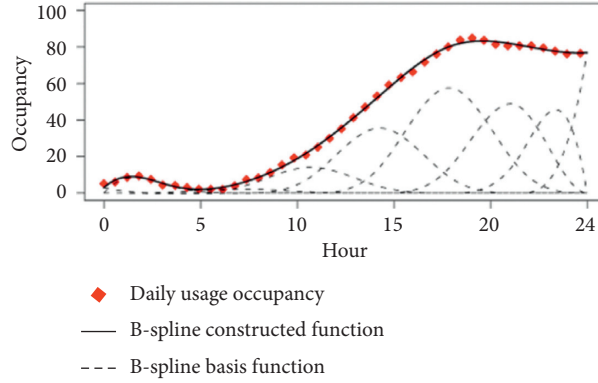


FIGURE 2: Illustration of B-Spline expansion.

functions were determined by the basis coefficients α_i^T , β_i^T , and γ_i^T . Such basis expansion method was advantageous in terms of transferring a high volume of data points into several basis functions' coefficients without losing the original pattern [11].

To obtain the basis coefficients α_i^T , β_i^T , and γ_i^T , the least square regression model was constructed as follows. $u_d(t)$ was used as an example to avoid repetition, but the method presented hereinafter was directly applicable to $e_d(t)$ and $u_j(t)$ as well.

$$u_SSE_i = \sum_{k=1}^{K_i} [u_d(t) - \alpha_i^T \varphi(t_{i,k})]^2. \quad (8)$$

To simplify the notations for the least square model, some matrix-formed data were introduced as follows:

$$u_d(t) = \begin{bmatrix} u_{di}(1) \\ u_{di}(2) \\ \dots \\ u_{di}(k) \end{bmatrix}, \quad (9)$$

$$\Phi_i = \begin{bmatrix} \varphi_1(t_{i,1}) & \varphi_2(t_{i,1}) & \dots & \varphi_M(t_{i,1}) \\ \varphi_1(t_{i,2}) & \varphi_2(t_{i,2}) & \dots & \varphi_M(t_{i,2}) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(t_{i,K_i}) & \varphi_2(t_{i,K_i}) & \dots & \varphi_M(t_{i,K_i}) \end{bmatrix},$$

where $u_{di}(k)$ was a vector that contained raw data points in day i . Φ_i was a $K_i \times M$ matrix; each column was the basis function value at all the time points. By reconstructing these usage occupancy data, the least square model can be rewritten in a simple quadratic form:

$$u_SSE_i = (u_{di} - \Phi_i \alpha_i)^T (u_{di} - \Phi_i \alpha_i). \quad (10)$$

Thus, the basis coefficients for day i can be estimated by

$$\hat{\alpha}_i = (\Phi_i^T \Phi_i)^{-1} \Phi_i^T u_{di}. \quad (11)$$

Through the B-spline model and least square regression, all three charging patterns defined above were converted into the basis coefficients. The functions can be obtained by $u_d(t) = \hat{\alpha}_i^T \varphi(t)$; $e_d(t) = \hat{\beta}_i^T \varphi(t)$, and $u_j(t) = \hat{\gamma}_i^T \varphi(t)$.

3.4. Functional Principal Component Analysis. After data smoothing, functional PCA was enabled as a powerful tool of the FDA approach to explore the curve's underlying features. In multivariate data analysis, PCA was commonly used to convert a large number of variables to some comprehensive variables that are much less in quantity but account for the highest variability. The mathematical solution of this problem was similar to finding the eigenvalue and the new variables were the functional principal components (FPCs).

In the FDA approach, the analyzed function contained information of a set of specific variables at enormous time points in a time interval. As a result, the work was confronted with the curse of dimensionality if the time was seen as the independent variable in the functional case. Consequently, the functional PCA method can be applied for the purpose of dimension reduction. In [19, 20], FPCA was employed as a data dimensionality reduction technique in the modeling of traffic flow patterns, which inhibit similar functional characteristics observed in EV charging. The approach was similar to the multivariate case. The dependent variable $x_i(s)$ ($s \in T$) was relative to x_{ij} in multivariate case.

$$f_i = \int \beta(s) x_i(s) ds = \int \beta x_i, \quad (12)$$

where $\beta(s)$ was the weight value and $\beta^k(s)$ denoted the weight function of k^{th} principal component. The variance function can be represented as $\text{Var}(f) = (1/N) \sum_{i=1}^N (\int \beta x_i)^2$.

Let f^k denote the k^{th} principal component, where $k = 1, 2, \dots, K$, so the relationship is $\text{Var}(f^1) > \text{Var}(f^2) > \dots > \text{Var}(f^K)$.

To calculate the first principal component, we just need to solve the following optimization problem:

$$\max \frac{1}{N} \sum_{i=1}^N \left(\int \beta^1 x_i \right)^2$$

s.t.

$$\int [\beta^1(s)]^2 ds = \int [\beta^1]^2 = \|\beta^1\|^2 = 1, \quad (13)$$

and the k^{th} principal component can be calculated by the following optimization problem:

$$\begin{aligned} & \max \frac{1}{N} \sum_{i=1}^N \left(\int \beta^k x_i \right)^2 \\ & \text{s.t.} \end{aligned} \quad (14)$$

$$\int [\beta^k]^2 = 1; \int \beta^k \beta^{k-m} = 0, \quad m = 1, \dots, k-1.$$

The covariance of $x(s)$ and $x(t)$ can be calculated by

$$v(s, t) = \frac{1}{N-1} \sum_{i=1}^N x_i(s)x_i(t). \quad (15)$$

The weight function of functional principal components $\beta(s)$ is needed to satisfy the following secular equation:

$$\int v(s, t)\beta(t)dt = \lambda\beta(s), \quad (16)$$

where λ was the eigenvalue and $\lambda_k / \sum_{k=1}^{N-1} \lambda_k$ meant the proportion of variability which the k^{th} principal component accounted for. The left side of (6) was an integral transform V of the weight function $\beta(s)$ with the kernel of the transform v defined by

$$V\beta(s) = \int v(s, t)\beta(t)dt. \quad (17)$$

The covariance operator was denoted by V . Therefore, (17) can be expressed as

$$V\beta(s) = \lambda\beta(s). \quad (18)$$

Equation (1) can be calculated through several methods, and we can calculate the FPC score f_i through (12).

4. Numeric Analysis

In this section, the numeric analysis was performed with the goal of understanding the EV owner's charging behavior patterns. The focuses were placed on three aspects: (1) variability analysis of the daily usage patterns of all EV charging stations, (2) variability analysis of the daily energy consumption of all EV charging stations, and (3) at the station level, the usage pattern differences analyzed.

4.1. Daily Usage Pattern Variability Analysis. To analyze the time-dependent usage pattern variabilities, the time-dependent occupancy of each day was calculated by aggregating all charging stations, so that in each year, a total of 365 curves were obtained, with each curve representing the occupancy of a day. Function PCA was then applied to extract the FPC from the dataset. It was observed that FPC1 accounted for 94% of the variance, and FPC2 accounted for 3%. When combined together, they reflected 97% of the data's variability and were kept for further analysis.

Figure 3 showed a way to look at the two FPCs and how they supported the unique analysis that FDA enabled. X-axis represented the time (0–24 hours in a day), and Y-axis represented the percentage of charging stations that were occupied at that time. The blue curve in both subfigures (a for FPC1 and b for FPC2) stood for the mean occupancy of all charging stations, while the green and red curves stood for the functions adding and subtracting one functional principal component. For example, in Figure 3(a), the green curve was generated by adding one FPC1 to the mean function represented by the blue curve, and the red curve was generated by subtracting one FPC1 to the mean function.

The first principal component focused on daytime between 7 am and 5 pm, which corresponded to the time that public charging stations were busiest in the day, especially workdays. Therefore, the first FPC essentially distinguished between working days and nonworking days. This observation was directly supported by Figure 4, in which almost all weekdays (blue dots) were located on the right-hand side of the plot, indicating a higher FPC1 score (X-axis), while almost all weekend days (red dots) were located on the left-hand side of the plot with lower FPC1 score. A few exceptions were identified in the plot and turned out to be the holidays, such as Labor Day and Independence Day, so these were nonworking days as well.

The second FPC accounted for only 3% variability and mainly captured the variance in the evening time from midnight to 6 am and again from 6 pm to midnight. The days with higher usage after 6 pm and before 6 am and with slightly less or average usage in the daytime would receive higher scores. However, due to the dominance of FPC1, the effect of FPC2 was rather limited.

Figure 5 presented the monthly and yearly charging usage patterns. The X-axis was the score of FPC1 and the Y-axis was that of FPC2. Figure 5(a) showed the monthly pattern with the colors standing for 12 months, respectively. No clear monthly pattern was observed.

Figure 5(b) showed the yearly pattern with the colors standing for years from 2014 to 2019. Dots in 2014 were almost invisible due to the low data size and overlap with pink color. The observation led to a clear pattern that as time went by, the scores of both FPC1 and FPC2 increased significantly. That meant that for the days with a higher FPC1 score, the occupancy continued to increase at an ever-increasing speed, while for the days with higher FPC2 scores, its morning and evening usage also increased significantly. This interpretation was in line with the rapid increase of EV ownership in Kansas City at a 78% year-over-year growth rate [21] and emphasized the needs for more charging infrastructures in the region.

Figure 6(a) shows the clustering result of the data. To make sure that similar data sizes are studied, data from 2016 to 2018 are selected for clustering. Compared with Figure 6(b), the result indicates that the data points in 2015

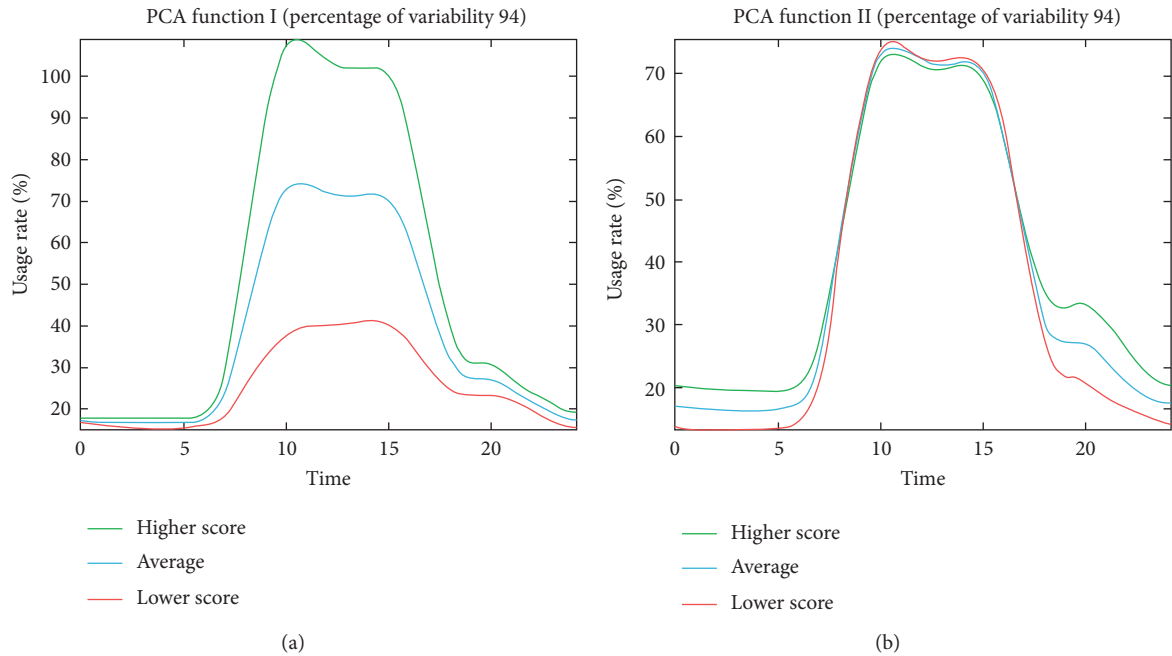


FIGURE 3: Effect of the first 2 FPCs on daily occupancy. (a) Left: the first FPC. (b) Right: the second FPC.

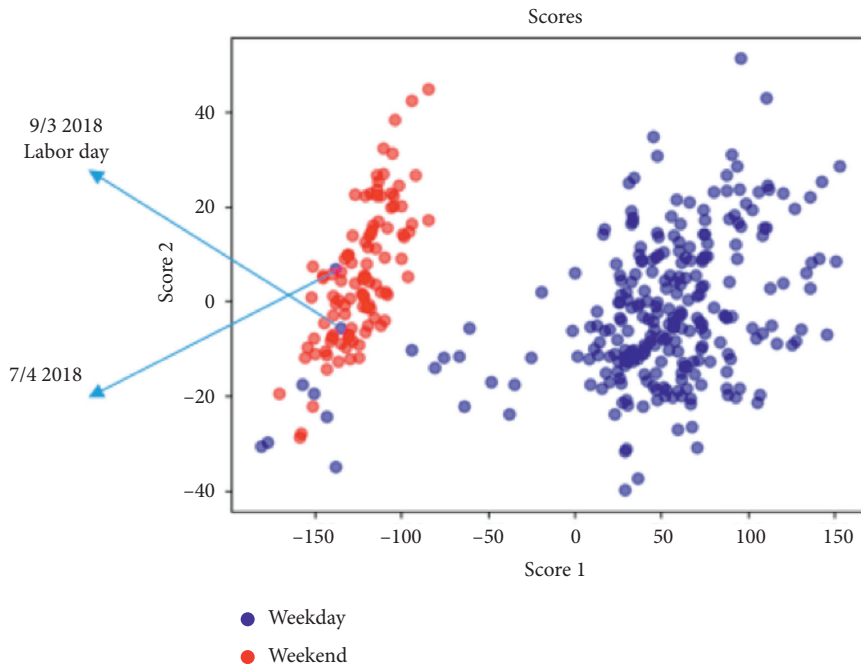


FIGURE 4: FPC scores of daily charging data.

have a lower FPC1 score and FPC 2 scores and are obviously separated from the other data points. However, the difference between 2017 and 2018 is less significant, which means that they have a similar occupancy pattern.

4.2. Daily Energy Consumption Variability Analysis. This section aimed to analyze the energy consumption variability caused by EV charging, which had a significant impact on the power grid and was helpful for grid load

management. Similar to the analysis above in Section 4.1, the time-dependent energy consumption of each day was calculated by aggregating all charging stations, so that in each year a total of 365 curves were obtained, with each curve representing the energy consumption of a day. Function PCA was then applied to extract the FPC from the dataset. It was observed that FPC1 accounted for 81% of the variance, and FPC2 accounted for 5%. So when combined together, they reflected 86% of the data's variability and

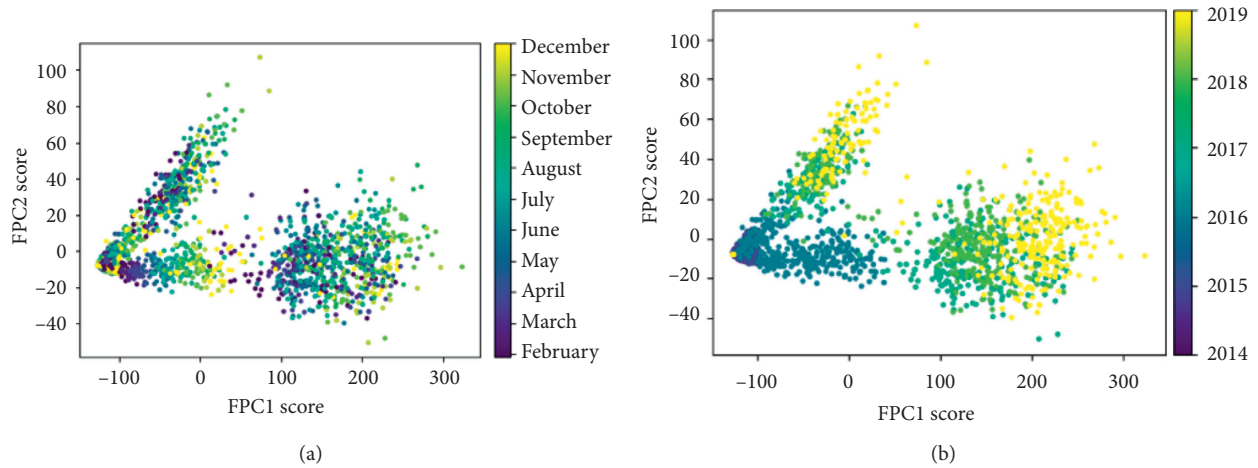


FIGURE 5: FPC scores for (a) monthly and (b) yearly pattern.

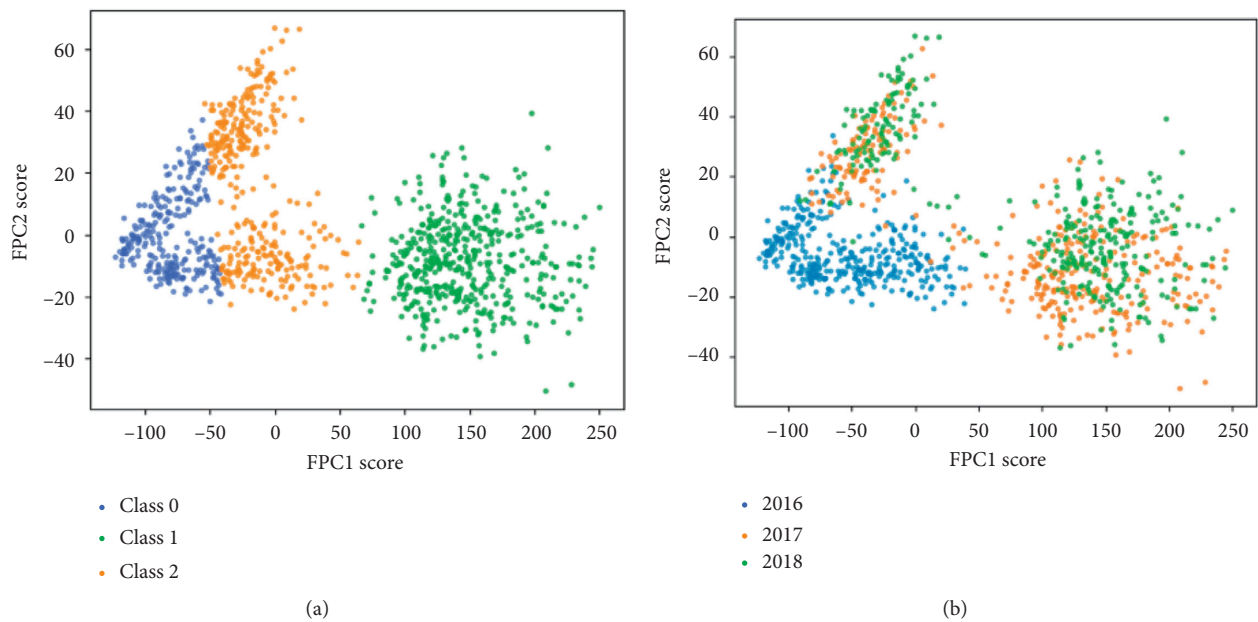


FIGURE 6: (a) FPC scores clustering (b) FPC scores during 2016–2018.

were kept for further analysis. The results were shown in Figure 6.

FPCA analysis of energy consumption resulted in some very different patterns when compared with daily occupancy. FPC1 mostly captured the variance of energy consumption during the morning peak between 7 am and 11 am. During this time range, the green curve increased dramatically, representing the days with a higher FPC1 score, and the required energy to charge EVs in the morning would be higher. On the other hand, FPC2 mostly captured the variance of energy consumption during the evening peak between 4 pm and 9 pm. In other words, if a day was observed to have a higher FPC2 score, its impacts on the power grid in the evening hours would be significantly increased.

A comparison between Figures 3 and 6 led to some interesting conclusions. While Figure 3 indicated that

from an occupancy perspective, the peak hour during the day started from as early as 7 am and did not end until 5 pm, Figure 6 suggested that the impact on the power grid became low after 11 am. This suggested that some vehicles did not leave the charging stations after they were fully charged, under which circumstances, the charging stations continued to be occupied (and thus unavailable to the other EV drivers), but from a power grid perspective, they did not require any energy. To validate such interpretation, the team went on to compare the charging event duration and the time EV actually spent on charging. The finding was as follows: while 40% of EVs left the charging stations within 1 minute after they are charged, the remaining 60% of EVs continued to park at the charging stations for various durations, and among them, two-thirds (or 40% of the entire population) would

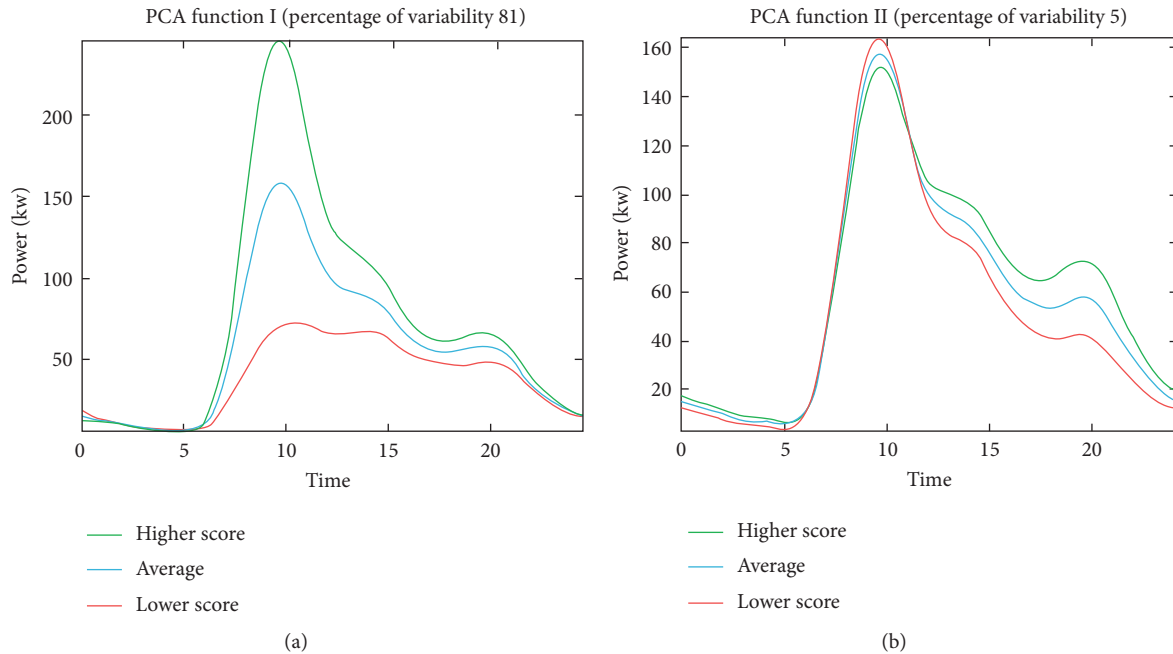


FIGURE 7: Effect of the first 2 FPCs on daily energy consumption. (a) Left: first FPC. (b) Right: second FPC.

even occupy the charging stations for at least an hour. While the discrepancy between EV owner's daily activity and the time needed for charging was understandable, the longer-than-reasonable parking behavior effectively reduced the availability of charging stations to the other EV drivers and, in our view, called for the need for parking regulation and enforcement.

4.3. Station Occupancy Variability Analysis. Different from the above analysis performed from the daily perspective, this analysis in this section examined the occupancy at the station level. So, each curve represented one charging station's 24-hour occupancy rates with all days aggregated, and a total of 455 curves (representing a total of 455 charging stations) were derived. Function PCA was then applied to extract the FPC from the dataset. It was observed that FPC1 accounted for 85% of the variance, and FPC2 accounted for 8%. So, when combined together, they reflected 93% of the data's variability and were kept for further analysis. The results were shown in Figure 7.

In the morning before 6 am and in the evening after 5 pm, stations with higher FPC1 scores were utilized more often than average, while in the daytime, their utilization rates were lower. On the other hand, stations with higher FPC2 values were utilized more often than average in the first half of a day (before noon) but were used less often in the second half of a day (afternoon).

An intuitive guess was these patterns might be attributed to the differences in the land-use patterns. As such, all 455 charging stations were mapped to five categories of land-use types: (1) recreational, which was

meant to be used for the enjoyment of the people who used it, such as arts center and theater; (2) commercial, which was designated for businesses, warehouses, shops, and any other infrastructures related to commerce, such as plazas, hotels, and hospitals; (3) transport, which was built for the structures that help people get from one destination to the other, such as airport; (4) industrial such as the plant and industrial parks; and (5) residential, such as apartments and condominiums. Then, the scores of FPC1 and FPC2 were plotted in Figure 8, in which Figure 8(a) had all land-use types together, while Figures 8(b)~8(f) stood for commercial, residential, industrial, transportation, and recreational.

No clear patterns can be found in Figure 9(a), in which charging stations of all land-use types were plotted together. However, when they were separated, some conclusions can be drawn. (1) For charging stations that were built on residential (Figure 9(c)), transport (Figure 9(e), and recreational (Figure 9(f)) areas, the majority of the dots in those subfigures had positive FPC1 scores. In other words, charging stations in these three categories shared a common pattern that they were used more often in the evening than in the daytime. Considering the nature of activities happening at these locations, this interpretation was consistent with what was observed in the real life. (2) FPC1 values for commercial (Figure 9(b)) and industrial (Figure 9(d)) areas were mixed and thus inconclusive to identify clear patterns. (3) Charging stations in the recreational area, in general, had negative FPC2 scores, meaning that they were used more in the second half of the day than before noon. Again, this seemed to be in line with our understanding of human behavior patterns.

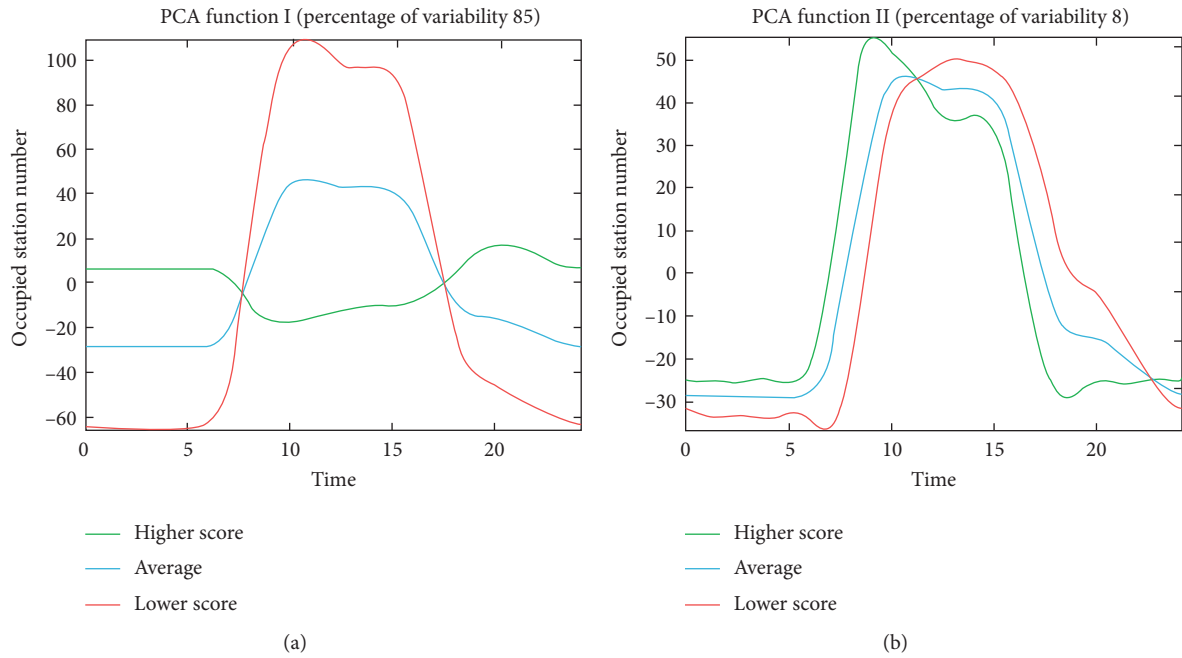


FIGURE 8: Effect of the first 2 FPCs on station occupancy. (a) Left: the first FPC. (b) Right: the second FPC.

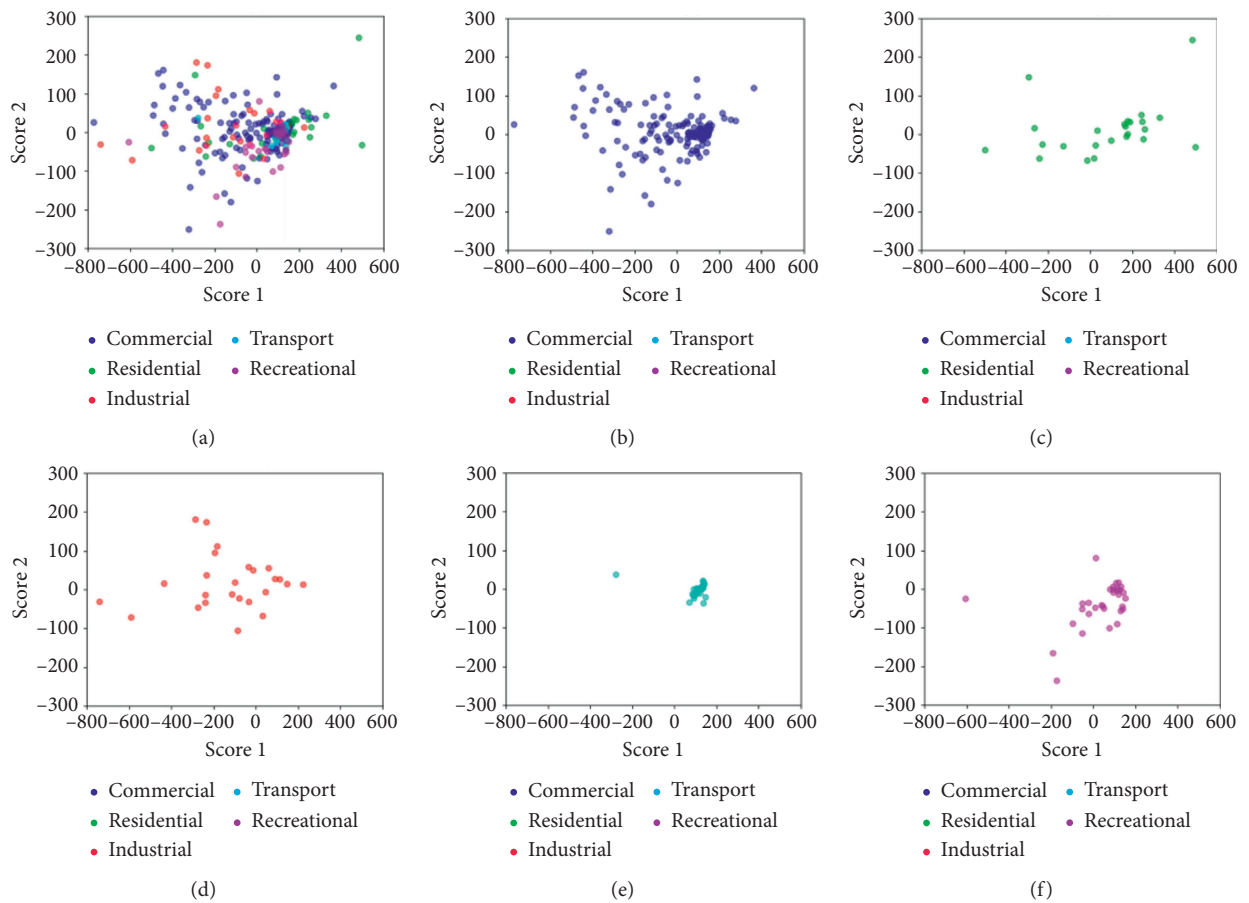


FIGURE 9: FPC scores of stations with different POI. (a) Scores. (b) Commercial. (c) Residential. (d) Industrial. (e) Transport. (f) Recreational.

5. Conclusions

In this manuscript, the focus was placed on analyzing the electric vehicle's usage behavior pattern with a functional data analysis approach, specifically, based on functional principal component analysis. Compared with the traditional discrete-based analysis models, the proposed FDA modeling approach had unique advantages in preserving the smooth function behavior of the data, bringing more flexibility in the modeling process with little required assumptions or background knowledge on independent variables, as well as the capability of handling time series data with different lengths or sizes. 5-year real-world charging event log data from a total of 455 charging stations in Kansas City, Missouri (KCMO), was used. The daily usage variability, daily energy consumption variability, and station-level usage variability were analyzed, with the goal of providing theoretical support to the EV infrastructure planning and regulation, as well as the power grid load management. In addition to the patterns revealed in the EV charging station's occupancy and energy consumption, the differences between EV driver's charging time and parking time were analyzed and called for the needs for parking regulation and enforcement. The different usage patterns associated with charging stations of different land-use types were also analyzed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This material is based on the work supported by the U.S. Department of Energy, Office of Energy Efficiency and

Renewable Energy's (EERE) Vehicle Technologies Office under the Award Number DE-EE008474. The authors are also thankful for the support from the Metropolitan Energy Center (MEC), the City of Kansas City Missouri (KCMO), Lilypad, Mid-America Regional Council (MARC), and Evergy (formerly Kansas City Power and Light Company (KCP & L)).

References

- [1] E. W. Wood, C. L. Rames, and M. Muratori, *New EVSE Analytical Tools/Models: Electric Vehicle Infrastructure Projection Tool (EVI-Pro)*, National Renewable Energy Lab.(NREL), Golden, CO, USA, 2018.
- [2] F. He, Y. Yin, and J. Zhou, "Deploying public charging stations for electric vehicles on urban road networks," *Transportation Research Part C: Emerging Technologies*, vol. 60, pp. 227–240, 2015.
- [3] J. He, H. Yang, T.-Q. Tang, and H.-J. Huang, "An optimal charging station location model with the consideration of electric vehicle's driving range," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 641–654, 2018.
- [4] O. Arslan and O. E. Karaşan, "A Benders decomposition approach for the charging station location problem with plug-in hybrid electric vehicles," *Transportation Research Part B: Methodological*, vol. 93, pp. 670–695, 2016.
- [5] Z.-H. Zhu, Z.-Y. Gao, J.-F. Zheng, and H.-M. Du, "Charging station location problem of plug-in electric vehicles," *Journal of Transport Geography*, vol. 52, pp. 11–22, 2016.
- [6] Q. Wu, A. H. Nielsen, J. Ostergaard et al., "Driving pattern analysis for electric vehicle (ev) grid integration study," in *Proceedings of the 2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, Gothenburg, Europe, October 2010.
- [7] P. Morrissey, P. Weldon, and M. O'Mahony, "Future standard and fast charging infrastructure planning: an analysis of electric vehicle charging behaviour," *Energy Policy*, vol. 89, pp. 257–270, 2016.
- [8] X.-H. Sun, T. Yamamoto, and T. Morikawa, "Charge timing choice behavior of battery electric vehicle users," *Transportation Research Part D: Transport and Environment*, vol. 37, pp. 97–107, 2015.
- [9] T. D. Chen, K. M. Kockelman, and M. Khan, "Locating electric vehicle charging stations: parking-based assignment method for seattle," *Transportation Research Record*, vol. 2385, no. 1, pp. 28–36, Washington, USA, 2013.
- [10] F. Jabeen, D. Olaru, B. Smith, T. Braunl, and S. Speidel, "Electric vehicle battery charging behaviour: findings from a driver survey," in *Proceedings of the Australasian Transport Research Forum*, Brisbane, Australia, October 2013.
- [11] X. Hu, Y. Yuan, X. Zhu, H. Yang, and K. Xie, "Behavioral responses to pre-planned road capacity reduction based on smartphone GPS trajectory data: a functional data analysis approach," *Journal of Intelligent Transportation Systems*, vol. 23, no. 2, pp. 133–143, 2019.
- [12] Z. Li, R. Kluger, X. Hu, Y.-J. Wu, and X. Zhu, "Reconstructing vehicle trajectories to support travel time estimation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 42, pp. 148–158, 2018.
- [13] X. Zhu, X. Hu, and Y.-C. Chiu, "Design of driving behavior pattern measurements using smartphone global positioning system data," *International Journal of Transportation Science and Technology*, vol. 2, no. 4, pp. 269–288, 2013.

- [14] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, CRC Press, Boca Raton, FL, USA, 1993.
- [15] S. Ullah and C. F. Finch, "Applications of functional data analysis: a systematic review," *BMC Medical Research Methodology*, vol. 13, no. 1, p. 43, 2013.
- [16] J. Ramsay and B. W. Silverman, *Functional Data Analysis*, Springer, Berlin, Germany, 2005.
- [17] X. Hu, X. Zhu, Y.-C. Chiu, and Q. Tang, "Will information and incentive affect traveler's day-to-day departure time decisions? - an empirical study of decision making evolution process," *International Journal of Sustainable Transportation*, vol. 14, no. 6, pp. 403–412, 2019.
- [18] X. Hu, X. Zhu, Y. Yuan, A. Rehan, Y.-C. Chiu, and M. Zmud, "A multi-resolution approach to investigate the impacts of pre-planned road capacity reduction based on smartphone gps trajectory data - a case study of mopac expressway in Austin, Texas," in *Proceedings of the Transportation Research Board 96th Annual Meeting*, Washington DC, USA, January 2017.
- [19] I. G. Guardiola, T. Leon, and F. Mallor, "A functional approach to monitor and recognize patterns of daily traffic profiles," *Transportation Research Part B: Methodological*, vol. 65, pp. 119–136, 2014.
- [20] I. M. Wagner-Muns, I. G. Guardiola, V. Samaranyake, and W. I. Kayani, "A functional data analysis approach to traffic volume forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 878–888, 2017.
- [21] Clean Charge Network, "Kansas city zooms ahead in electric vehicle growth," 2017, <https://cleanchargenetwork.com/kansas-city-zooms-ahead-in-electric-vehicle-growth/>.

Research Article

Development of a Model for Evaluating the Coverage Area of Transit Center Using Smart Card Data

Hyunjun Hwang ^{1,2} Shin-Hyung Cho ^{3,4} Dong-Kyu Kim ^{1,4}
and Seung-Young Kho ^{1,4}

¹Department of Civil and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

²Department of Transportation System Research, The Seoul Institute, Seoul 06756, Republic of Korea

³School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

⁴Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

Correspondence should be addressed to Seung-Young Kho; sykho@snu.ac.kr

Received 23 April 2020; Revised 18 October 2020; Accepted 29 October 2020; Published 19 November 2020

Academic Editor: Hong Yang

Copyright © 2020 Hyunjun Hwang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since metropolitan cities are broadening as a result of urban sprawl, multimodal transportation systems have been adopted to fulfill the connection between the suburban and urban areas. The transportation system is being revamped around the transit center in the urban area to facilitate access to the downtown area from the suburbs. Studies are being conducted to improve the accessibility of public transportation by using the concept of hub-and-spoke. In this study, we develop a coverage area index (CAI) to assess the impact of a transit center on access to urban areas from the suburbs quantitatively. The concept of network centrality and the kernel density function is used to evaluate the extent of the influence of a transit center. The smart card data in the Seoul metropolitan area are used to analyze the CAI. Six transit centers in the Seoul metropolitan area are investigated to compare the coverage area to the transit center. The bandwidth of the kernel density function is set as 2 km considering the size and influence of each region. We evaluate six transit centers using the CAIs in Seoul compared to the index characteristics with transit accessibility (TA) index from previous studies. The CAI is possible to identify the incompetent centers, alternative routes to solve the problems of overcrowding on the centers, and areas with insufficient supplies of regional transit.

1. Introduction

As the population increases, urban sprawl can occur rapidly, and it can impinge on the suburban areas. A hub-and-spoke structure of the integrated multimodal public transit system is adequate for covering large suburban areas with restricted infrastructure and operation costs. Travelers who wish to go downtown take regional express transit from suburban areas to access a transit center near the downtown area, where they can transfer to a city bus or an urban railway to take them downtown. This system is cost-effective, but passengers may be uncomfortable due to transfers if transit modes are not well coordinated.

Transit centers serve as coordination points between downtown and the suburban areas in a hub-and-spoke structure

network. To evaluate the functionality of transit centers, operators should investigate the area influenced by the centers. The investigation will provide operators with information concerning the expected boundary of influence when some actions are implemented at the transit centers. Conversely, poor connectivity of a suburban area to the transit centers means that trips from the suburban area to downtown are inefficient.

Several studies have evaluated the effectiveness of transit centers using a data envelope analysis (DEA) model [1, 2]. Some studies have developed an index of connectivity and accessibility with variables like headway and transfer time [3, 4]. In the planning phase, other studies have used accessibility and connectivity indices derived from the socioeconomic indices of neighborhoods to evaluate transit centers [5, 6], and they have used heuristic models to

evaluate the locations of transit centers [7]. However, such models have limitations that they do not consider the origins and destinations of transferred travelers. To overcome this limitation, we suggested developing an index using smart card data to know the origin and destination of transit center users.

This study aims to develop a coverage area index (CAI) for transit centers considering the center's role which is to facilitate the connection in the hub-and-spoke structure of the regional transit system that connects the suburban areas with the downtown area. The concept of betweenness centrality is applied to the origin and destination of actual trips, aiming to indicate the travelers' preferences. The CAI is defined as the proportion of the number of trips through a transit center to the number of total trips taken from the suburban area to downtown. The kernel density function is employed to aggregate discrete trip chain data to traffic analysis zones. As the index usage, we discuss a comparison of transit centers, transit supply, and structure of the transit network.

Research on the assessment of transit centers includes relative evaluation between centers with empirical data, defining and measuring indices, using socioeconomic factors to identify optimal hub locations during the planning phase and developing indices using a network centrality concept from social science. There have been several studies on the data-based evaluation of transit centers using the DEA model. The relative efficiency of individual centers was calculated through mutual comparisons of the centers' defined inputs and outputs. DEA model was employed with the inputs, e.g., the areas of the hubs, the cost of operation, the number of staff members required, and capacities of the hubs. The model's outputs were the number of transfers per day, average, and deviations of the transfer times [1]. Another research employed DEA model with the inputs, e.g., transfer time, number of subway and bus lines, and waiting time of the hubs and outputs of the model were number of transfer trips and transfer rate [8]. In Kochi City, Japan, the same DEA model was used with smart card data to evaluate transfers between buses and trams. Headways and the number of lines were used as input, and the number of transfers and the time required for transfers was the desired outputs of the model [2]. The DEA model was useful for assessing the efficiency of the transit centers, but it was difficult to evaluate them in conjunction with the surrounding area.

Accessibility and connectivity were common indices used to evaluate the transit centers, and there were various detailed definitions in each concept. From the first perspective, connectivity was regarded as an index for evaluating the availability of transit. A study evaluated the connectivity of nodes, lines, transfer centers, and networks using the connecting power calculated from the headway and capacity of the transit lines [3]. From the second perspective, connectivity was regarded as an index for evaluating travelers' experiences. There was a study that defined connectivity as the transfer time between modes and that measured the index with empirical data from a transit corridor in Auckland, New Zealand [4].

Both views treated connectivity as a function of capacity and time, but not considered the spatial distribution of trips.

Accessibility of the transit facility was evaluated based on the population or the number of jobs within a small distance from the facility [9, 10]. In the Transit Capacity and Quality of Service Manual (TCQSM), the traffic analysis zones' accessibility index was defined as the percentage of an area within a certain accessing distance of stations among the transit-supportive area [11]. A study measured the accessibility of zones with a weighted average of several indices to find services gaps [12]. Accessing distance and coverage of a facility were often considered as a walking distance to access the station, but some studies expanded the concepts of accessing distance and coverage to other modes. One study measured the accessibility of a station with a total area that could be reached in a reasonable time by walking, riding a bicycle, and using local bus services [13]. Another study showed a transit center's access zone with proper access time to reach the center with multimodal transit [14]. Another accessibility measure that considers the origin-destination (OD) of trips was motivated by the gravity model of demand forecasting [15]. There have been many studies that measured accessibility using the supply of transit service, distance, and time. In contrast, there have been only a few studies that focused on the transit center and the actual OD pair.

At the planning level, the locations of transit centers were optimized to maximize the network's performance considering the center of mass as the optimal location of the hub. This approach had a limit in that it assumed that distance was the only variable in the relationship between a transit center and the coverage area [5]. In another approach, it was assumed that the accessibility index is the neighborhood population not overlapping with other neighborhoods. The connectivity index was viewed as the number of connections between nodes and areas with large numbers of travelers to optimize the overall transit network [6].

Some approaches have used the network centrality concept from network theory to evaluate transportation networks. Several centrality measures were used to evaluate traffic analysis zones in Seoul, Korea, with a kernel density function to aggregate network nodes to zones [16]. Specifically, closeness centrality was used to represent the accessibility of transit nodes [17].

Studies used smart card data to assess the accessibility of transit and acquire total ridership of the transit network. There was a study that developed a platform using smart card data and vehicle GPS data. The platform was designed to visualize transit performance measures in several levels—network, route, and stop level [18]. Another approach of using smart card data was finding a relationship between the data and demographics, land use, and transportation factors to predict ridership or purposes of trips [19, 20]. Various data other than ridership can be obtained from smart card data in Seoul where the departure and destination time and location were specified. There were studies used the travel time obtained from the data to assess the competitiveness of transit network compared to other

modes [21, 22]. There was another approach using mobile phone data to get the origin and destination of trips with unknown modes. The origins and destinations, travel time, and distance of the private car and public transit were compared to define the accessibility index [23]. These studies presented some ways to assess accessibility by data-driven approach, but they did not use the origin and destination of transit trips. Evaluating transit centers, not only origin and destination but also the transfer locations of transit trips, also needed to be considered.

As indicated above, there are several ways to evaluate transit centers, but there is no specific measure that can be used to evaluate the pivotal role of transit centers as connecting nodes between suburban areas and the downtown area with data from actual trips. Thus, if we understand the connections in the transit center, our understanding of the overall distribution of trips in the metropolitan area can be improved. At the network level, centrality can be one of the concepts that address connecting the role of a transit center, but previous research does not concentrate on transit centers.

2. Model Formulation

2.1. Definition of Coverage Area for a Transit Center. The transit center was defined as a meeting point of different modes and several routes and lines that function as effective hubs working as a regional collection-distribution function [7, 11]. The original concept of coverage was that “something can be understood as an area in which the influence of such a thing can be perceived” [24]. The coverage of public transportation facilities refers to the range of space accessible by walking or riding a bicycle. For a rail station in the suburban area, the coverage can be extended to an accessible zone with automobile access, such as drop off or kiss-and-ride locations in a particular case [11]. In this study, because the transit center is the location where many travelers transfer from fast regional transit to city transit, the access modes to stipulate the coverage of transit centers were expanded to every mode, including regional buses. From this perspective, the coverage area for each transit center is the set of suburban zones which are easily accessible and strongly dependent on the transit center to approach the central city area. The expanded concept of a coverage area in a metropolitan area is shown in Figure 1. Travelers departing from the suburban zones go through transit centers in the route of a trip to the central city. Thus, with a well-constructed hub-and-spoke structure, a large portion of travelers transfers from the regional transit to other modes at the transit centers. This phenomenon is similar to the “betweenness” concept in network theory. To analyze the expanded coverage, we developed a coverage area index using betweenness centrality concepts. OD trips were used to measure centrality instead of a transit network topology because the topology was too complicated to reflect travelers’ preferences. The kernel density function was used to obtain correctly aggregated OD trip data from smart card data.

2.2. Network Centrality. Network centrality is a measure to evaluate the nodes of networks in social science and network theory. There are several kinds of centrality, e.g., betweenness centrality, closeness centrality, and straightness centrality [25–27]. Betweenness centrality is defined as how a zone is located between all pairs of other zones. It is measured as the ratio of the number of shortest paths through the zone to the total number of the shortest paths between all pairs of zones. It is a suitable centrality measure to evaluate the coverage area of transit centers because these centers’ role in transit network is the connection node between zones in the coverage area and the downtown zones. Betweenness centrality can be expressed mathematically as follows [16]:

$$C_B(i) = \frac{1}{(N-1)(N-2)} \sum_{\forall j,k | j \neq k} \frac{n_{jk}(i)}{n_{jk}}, \quad (1)$$

where $C_B(i)$ means betweenness centrality of zone i , N means the number of zones, $n_{jk}(i)$ means the number of the shortest paths from zone j to zone k through zone i , and n_{jk} means the number of the shortest path from zone j to zone k .

To apply the definition for the coverage area, trip chain data are used instead of the shortest-path data. The term of “number of the shortest paths” is substituted with the number of trips due to traveler choice behavior to define CAI.

2.3. Kernel Density Function (KDF). When aggregating actual trip data, boarded or alighted zones recorded in the smart card data are not the actual origin or destination zones of the trips because travelers walked or took other modes without records to the neighbor zones to access transit. The kernel density function (KDF) is used to apply this neighborhood effect. The value of KDF is set by the type of function and the bandwidth. In this paper, we use the simple triangular kernel function. The influence of neighbor areas decreases linearly with distance, and when the distance exceeds the bandwidth, the influence became zero. KDF can be formulated as

$$K(d_{ik} | w) = \begin{cases} 1 - \frac{d_{ik}}{w}, & \text{if } d_{ik} < w, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where d_{ik} means the Euclidean distance between the centroids of zone i and k (km) and w means the kernel bandwidth (km).

The bandwidth of KDF is the only variable that affects the power of the neighborhood effect, but the optimal bandwidth is not defined. In this research, we set the bandwidth to 2 km, which is the median of the Euclidean distance between adjacent zones in the Seoul metropolitan area. This bandwidth is justified based on the case study.

2.4. Coverage Area Index (CAI). Inevitably, suburban zones in a metropolitan area can be connected directly to only some points of the central city because of the limitation in

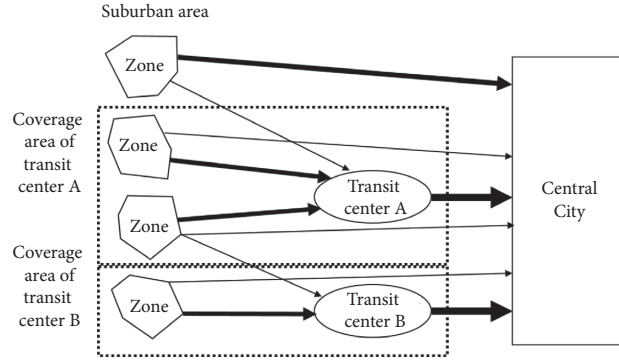


FIGURE 1: An expanded concept of a coverage area of the transit center.

the infrastructure or the running cost. Given this constraint, travelers have to transfer from regional transit to city transit, such as an urban bus or the metro, and transit network should be formed as a hub-and-spoke structure. Existing accessibility indices for a transit center have rarely been focused on the complexity of the transit network and travelers' choice behaviors caused by the structure of the network. Thus, this study uses actual trip data that reflects all travelers' personal preferences rather than using the shortest path in betweenness centrality. Actual trip data allow us to analyze the dependence of travelers from each suburban zone on specific transit points by using the definition of betweenness centrality inversely. Since actual trip data represent current travelers' travel behavior, it would be suitable to analyze the current coverage rather than predict the future or potential coverage.

From equation (1), the central city was considered as zone j , and each zone in the suburban area was considered as zone k . The number of shortest paths between zones j and k was transformed into the number of trips between the central city and each zone. The number of shortest paths through a transit center, i , was transformed to the number of trips between the central city and the zone that included a transfer at the transit center. This was reasonable because a route that is chosen for a trip indicates that the route had the smallest general cost among the possible choice set.

The coverage area index of zone i for transit center j (CAI_{ij}) is defined as the ratio of the modified number of trips through the transit center j to the total number of trips taken from zone i to the central city. KDF is used to account for the neighborhood effect and modify the original number of trips. The definition of CAI is

$$CAI_{ij} = \frac{T'_{ijC} + T'_{Cji}}{O'_{iC} + D'_{iC}}, \quad (3)$$

$$\begin{aligned} O'_{iC} &= O_{iC} + \sum_{k \neq i} K(d_{ik}|w)O_{kC}, \\ D'_{iC} &= D_{iC} + \sum_{k \neq i} K(d_{ik}|w)D_{kC}, \\ T'_{ijC} &= T_{ijC} + \sum_{k \neq i} K(d_{ik}|w)T_{kjC}, \end{aligned} \quad (4)$$

where CAI_{ij} means the coverage area index of zone i for transit center j , O_{iC} means the total number of trips to the central city from zone i , D_{iC} means the total number of trips from the central city to zone i , d_{ik} means the Euclidian distance between the centroid of zones i and k , T_{ijC} means the number of trips from zone i to the central city through transit center j , T_{Cji} means the number of trips from the central city to zone i through transit center j , and $K(d_{ik}|w)$ means the kernel density function with distance d_{ik} and bandwidth w .

CAI is the ratio, so it has a value between 0 and 1. When the value is close to 1, most travelers from the zone to the central city transfer to city transit at the center. This indicates that the zone is in the area covered by the transit center. When the index is close to 0, almost none of trips from the zone to the central city passed through the transit center. If the central city is big enough to have more than one transit center, the index can be measured individually for each transit center.

We employ the two methods for network evaluation, which are network centrality and kernel density function. To evaluate the transit center, betweenness centrality is used to define the geographical location from other zones, and kernel density function is applied to catch the effective area from the transit center. While betweenness centrality is focused on one zone or node, CAI considers the trips between transit center and destination compared to other transit centers. It is shown that comparing equations (1) and (3), betweenness centrality is defined as $C_B(i)$ and CAI is defined as CAI_{ij} . This characteristic gives CAI a unique usage that not only evaluates transit centers but also tests transit network of each zone with spatial distribution of CAI.

2.5. Evaluation of CAI Compared to Accessibility Index.

The CAI usage was evaluated by the contrast of the CAI and the accessibility measure presented in the previous study. The accessibility measure used in the study was defined as the product of attractiveness at destinations and friction between zones and destinations. This interpretation came from the demand forecasting model with socioeconomic conditions and transit networks [15], using parameters and

variables from Sacramento, CA [28]. Because this study was focused on Seoul, Korea, to reflect local characteristics, the parameters and variables from Korea's national demand forecasting model were used [29]. The model was derived

from a national survey and used a gravity model to forecast. To calculate TA with the model, some of the variables needed to be modified while maintaining the TA's fundamental meaning. The definition of TA is

$$\begin{aligned}
 TA_{ij} &= (CATT_j) * (ZFRIC_{ij}), \\
 CATT_j &= 2.296 * EMP_j + 0.536 * POP_j + 0.443 * EMPS_j + 0.801 * STU_j, \\
 ZFRIC_{ij} &= \begin{cases} 0.1093 * MTD_{ij}^{1.58808} * \exp(-0.2437 * MTD_{ij}), & \text{if } MTD_{ij} < 13.0 \text{ km,} \\ 2.4043 * MTD_{ij}^{-0.73839} * \exp(-0.0190 * MTD_{ij}), & \text{if } MTD_{ij} > 13.0 \text{ km,} \end{cases}
 \end{aligned} \tag{5}$$

where TA_{ij} means the transit accessibility between zone i and center j ; $CATT_j$ means the attraction of the transit center j ; $ZFRIC_{ij}$ means the friction of access between zone i and the center j , which was defined as a combined model; EMP_j means the total number of employers in the zone of transit center j (people); POP_j means the population of the zone of transit center j (people); $EMPS_j$ means the total number of people working in service in the zone of transit center j (people); STU_j means the total number of students in the zone of transit center j (people); and MTD_{ij} means median travel distance by transit between zones i and j .

3. Case Study

3.1. Data Description. Smart card data in the Seoul metropolitan area were used to obtain public transit trip chain data, and the data covered almost 99% of trips by public transit in Seoul. Travelers' boarding time and station, arrival time and station, modes used, the number of transfers, a unique ID for each smart card, and a transaction ID for each trip chain are recorded because users have to tag their card both when they get on and off transit modes in Seoul. With appropriate treatment using the transaction ID recorded in the data, all trip chains between the suburban area and Seoul and locations they passed through can be obtained. In this research, 180 million trip data for five weekdays (May 16–19 and May 22, 2017) were used. Data for 2.5 million trips (2.33% of total trips) were not used because they did not include arrival time and station. After treatment, 15.77 million trip chains were obtained between suburban areas and Seoul.

Six transit centers were selected for the case study. They serve as the destinations for most of the regional express buses from the Seoul suburban area because they are accessible to the highway, and they are connected to the metro lines in Seoul. Three of them (Gangnam, Sadang, and Jamsil) are in the southern part of Seoul near the Gangnam business district, subcenters of Seoul. The other three (Seoul Station, Cheongnyangni, and Hapjeong) are near the central business district, the historical downtown of Seoul, in Figure 2.

About two million travelers transferred their modes from regional bus to the city bus and subway or vice versa at the centers each day. An average number of trips between each zone in the suburban area and Seoul was 4,526.5 per

day, and 13.8% of the total trips (626.8 trips per day for each zone) passed through the transit centers. According to the percentage of trips, it could be shown that the transit centers were prominent in the regional transit network of Seoul. The average travel time and the travel distance of trips through the transit centers were 49.4 minutes and 26.9 km. The average travel speed of trips through the transit centers was faster than the speed of total trips between the suburban area and Seoul. The basic statistics of the result of the data treatment is shown in Table 1.

3.2. Effect of the Bandwidth of KDF on CAI. Stations often were located near the boundaries of the zones, and such stations only counted as one zone even though people from neighboring zones used them. In addition to this, smart card data might not reflect the actual origin of the trip because people often cross the neighboring zone of their origin to access transit modes. This effect was adjusted by applying the neighborhood effect with KDF when aggregating discrete smart card data to spatial zones.

The median distance between centroids of neighbor zones in the study area is 2.1 km, and the average zone size is 16 km². We assumed that people would not cross more than one zone to access transit. Sensitivity analysis on the bandwidth was conducted based on value of 2 km. The bandwidth of 0 km—without KDF, 2 km, and 5 km was used. The measured result of CAI of Gangnam in part of the region is shown in Figure 3. Without applying KDF, some adjacent zones had significant differences in CAI. When the bandwidth of the kernel function was 5 km, the index flattened in a large area, and the power of adjacent zones increased. This resulted in smaller differences in the index between adjacent zones.

According to the sensitivity analysis, the proper bandwidth of KDF depends on the purpose of the evaluation. Short bandwidth is suitable for observing the differences in a small view, such as the zone level. Conversely, long bandwidth is suitable for observing differences in a broad view, such as the regional or in-county levels. In this study, we used 2 km for the bandwidth to observe a difference between adjacent zones because we wanted to compare the difference in CAI in zone level.

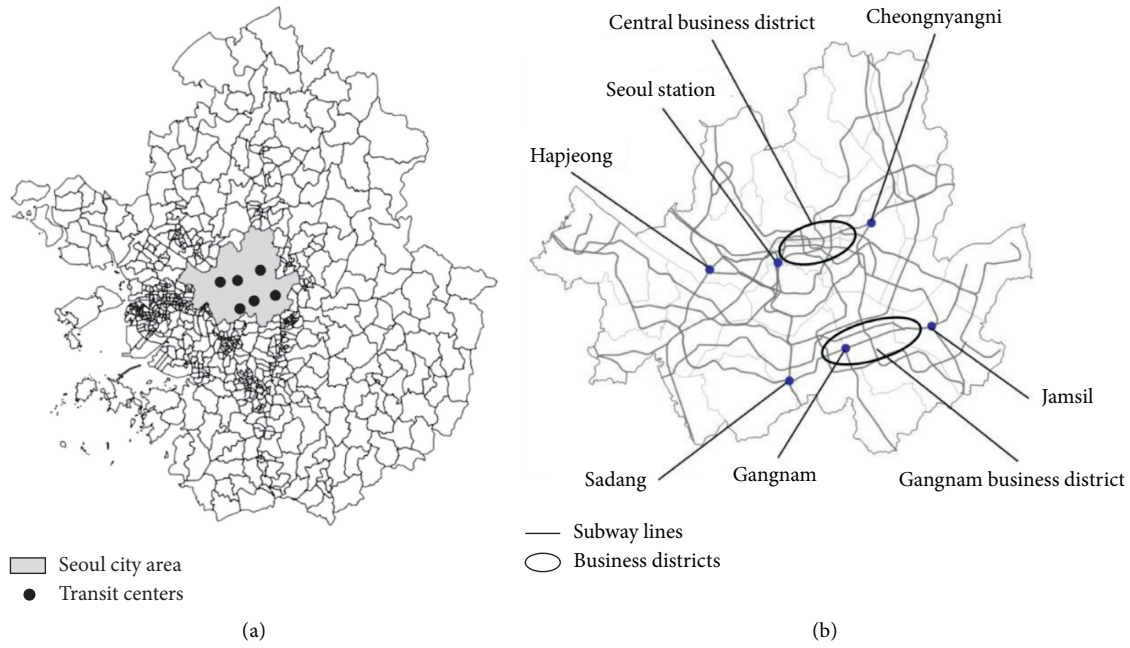


FIGURE 2: Study area: (a) outline of the Seoul metropolitan area; (b) transit centers used in the analysis.

TABLE 1: Basic statistics of the obtained trip chain data between suburban zones and Seoul.

Division	Total # of trips	# of regional bus lines	Travel time (min)	Distance (km)	Travel speed (km/h)
Total (697 zones)	15,775,018		51.4	23.0	26.8
Trips through the transit centers	2,184,397		49.4	26.9	32.7
Gangnam	741,220	59	50.9	30.3	35.7
Jamsil	389,536	35	45.4	23.3	30.8
Sadang	442,237	15	39.7	23.0	34.8
Seoul Station	367,146	34	58.7	29.6	30.3
Cheongnyangni	114,217	17	63.6	26.6	25.1
Hapjeong	130,041	18	47.8	24.4	30.6

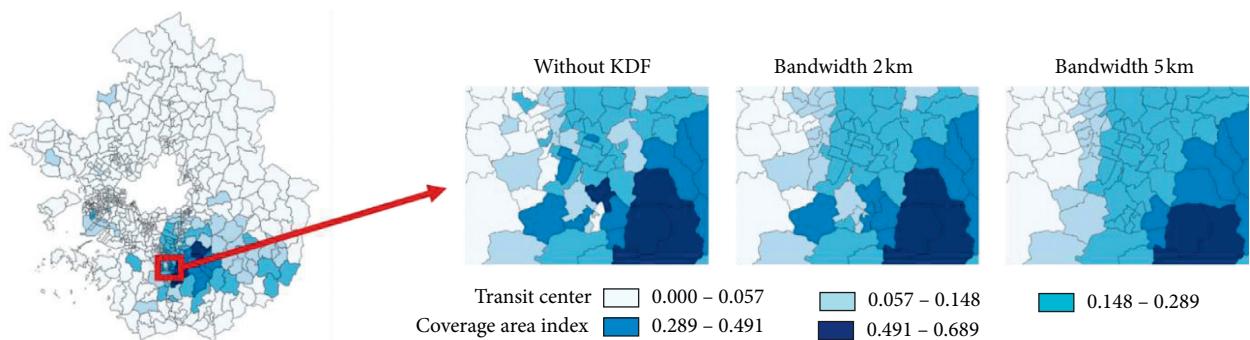


FIGURE 3: Effect of kernel bandwidth to CAI of Gangnam in the southern part of the Seoul suburban area.

3.3. Coverage Area of Transit Center Using CAI. The overall coverage area of each transit center is shown in Figure 4. The coverage area was defined as the zone in which the biggest CAI value was bigger than 0.05 and had more than 100 trips to Seoul. Sadang covered the area in the southwest part of the Seoul suburban area. The CAI for Sadang had the highest value of 0.922 at Jangan. This indicates that 92.2% of the trips from the Jangan zone to Seoul passed through the Sadang transit center. The Gangnam transit center is close to the

most important highway in Seoul, and it is the transit center that is closest to the Gangnam district. These characteristics of the location led to the broad area that it covers. Cheongnyangni and Jamsil are located in the eastern part of Seoul, so their coverage areas are shown in the eastern part of suburban Seoul. Seoul Station and Hapjeong covered the west part of Seoul suburban area, but CAIs of each zone to them were smaller than the value of other centers. This tendency means that Seoul Station and Hapjeong were not

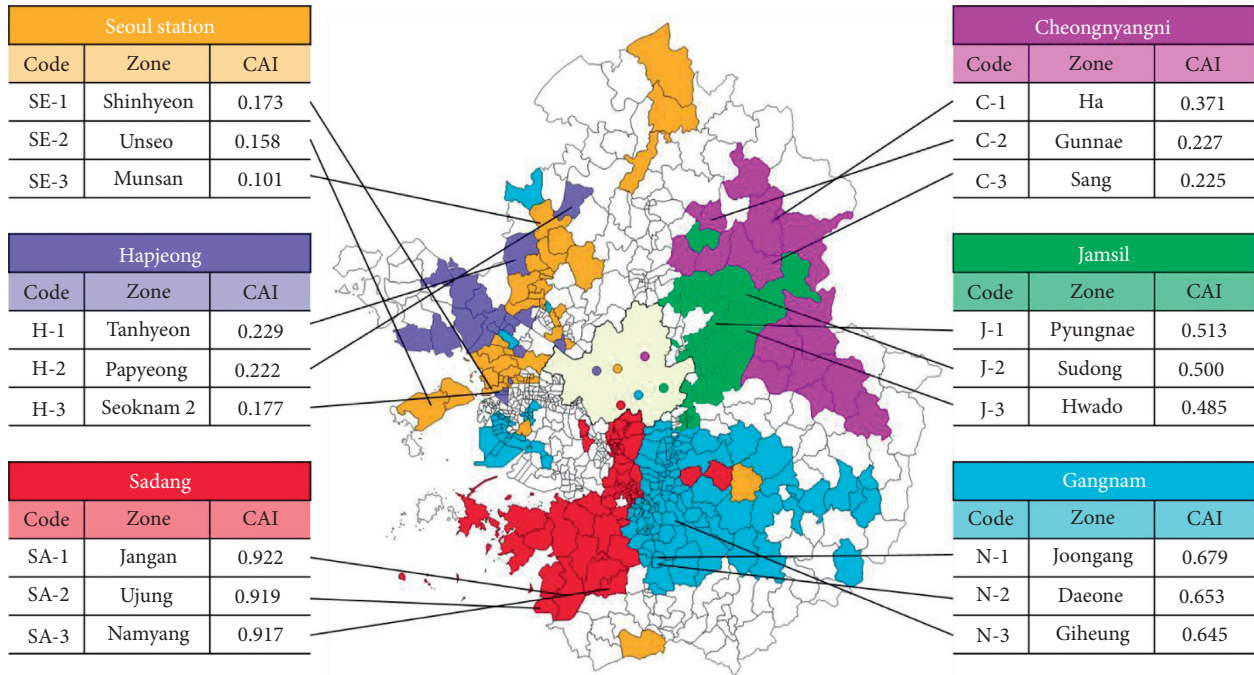


FIGURE 4: Areas influenced by each center and the three regions in each center that have the highest indices.

functional as other centers and less important than other centers.

Seoul Station and Gangnam show the most significant difference in the public transportation network. As shown in Figure 5, they had a similar distribution of TA, and they both located in major business districts of Seoul, but their CAI was very different from the centers. Gangnam had a highly concentrated CAI in the southern part of Seoul. It means that the transit network in the southern part of Seoul had a hub-and-spoke structure centering on Gangnam. In contrast, the coverage of Seoul Station located in both southern and northwestern parts of Seoul and CAI values of the area was low. It means that Seoul Station was connected to the board part of the Seoul metropolitan area, but the power of the center was not strong enough to constitute the hub-and-spoke structure transit network. Through this analysis, we could discuss that transit lines connecting Seoul Station and the southern part of Seoul should be adjusted to connecting the center and northeastern part of Seoul to achieve the effective hub-and-spoke structure of transit network centering Seoul Station. In terms of transit in downtown, it is necessary to examine whether travelers arriving at Seoul Station from a suburban area can travel to other parts of the city adequately.

Consequently, CAI can identify transit centers with not enough power as a hub in a hub-and-spoke structure; it means the functionality of the transit center in the overall network would not be sufficient. With Figure 4, three centers with a lower CAI over their own coverage—Seoul Station, Hapjeong, Cheongnyangni—would correspond to the case. The centers should be upgraded to be attractive to travelers, such as speeding up regional buses with bus rapid transit or

making a convenient environment to transfer with a high-quality in-door bus shelter and information system.

From a broad perspective rather than an individual center, the coverage zones' CAI of the transit centers in the central business district (Seoul Station, Hapjeong, and Cheongnyangni) were smaller than the index of the centers in Gangnam district (Sadang, Jamsil, and Gangnam). We can infer that the transit network near the central business district was not forming a hub-and-spoke structure. Therefore, to make the transit network of the district efficient, it can be considered that the government should reform the transit network or transit centers of the district to forming a hub-and-spoke structure.

3.4. Evaluation of Index Performance. We compared CAI and TA for three centers—Sadang, Gangnam, and Seoul Station—of total six centers. Sadang was a hub that performs well, and it had a well-defined coverage. Gangnam and Seoul Station were located in two different subcenters of Seoul. They had a similar distribution of TA, but their CAI shows the difference between the two hubs. The trend shown in Figure 5 indicated a significant difference between CAI and TA. There were some core zones in the suburban area with high CAI values, but CAI decreased rapidly when the distance from the core increased. TA was shown to have its highest value near the central city, and it decreased slowly as the distance from the central city increased. The reason for this difference in trends was that the friction of the TA index was affected strongly by the Euclidian distance from the transit center. In contrast, CAI was not strongly affected by the distance between the zone and the central city. Due to

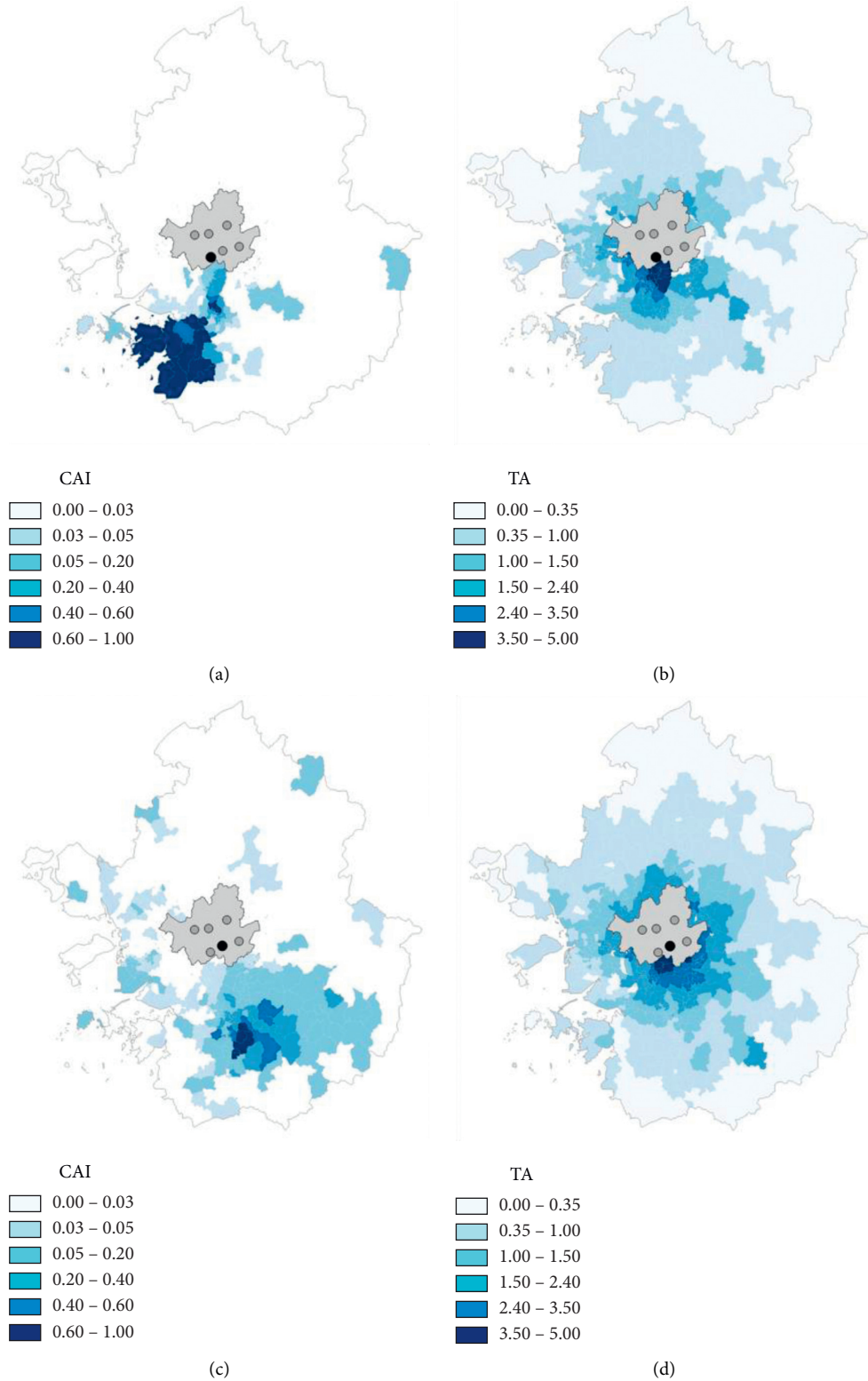


FIGURE 5: Continued.

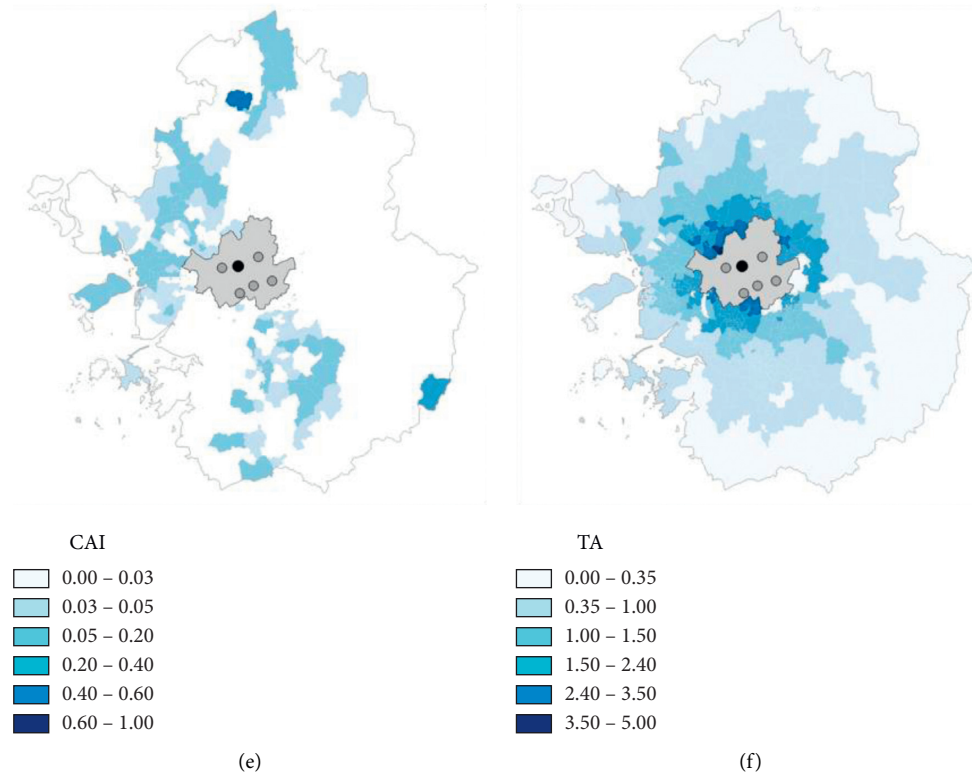


FIGURE 5: Comparison between CAI and TA of three transit centers: (a) CAI of Sadang; (b) TA of Sadang; (c) CAI of Gangnam; (d) TA of Gangnam; (e) CAI of Seoul Station; (f) TA of Seoul Station.

CAI's definition, the ratio of the number of trips was not affected directly by the distance.

This result showed the contribution of CAI. Almost all zones near the central city are connected to city transit. Travelers from those zones do not have to transfer at the transit centers. Zones connected to a transit center with a regional express bus line would have high CAI to the center and would be affected significantly by the center's renovation even though they are far from the city. If the bus line were terminated, the level of service would decrease dramatically. This kind of spatial dependency analysis can offer new motives and insights into planners and operators. The CAI and other accessibility indices, such as TA, have complementary relationships.

4. Discussion

4.1. Measure of Transit Supply and User Preference. A zone with high CAI of a transit center means a high percentage of travelers between the zone and the central city was integrated at the center. Except for the case where the total number of trips was small, a route passing through the center was the route most preferred by travelers when they chose a route between the zone and the city. This means that the zone is connected to the center with a high level of regional services, such as an express bus line in the case of Seoul. If a zone has a small CAI of a transit center, travelers from the zone to the city will transfer to city transit at another center

or out of the city. The latter case implements a longer trip by slower transit modes, such as the metro or a city bus.

Observing the result of CAI with the existing regional transit network, especially the main corridor of express bus lines, it was apparent why the index is what it is. Figure 6 shows that the zones with high CAI were connected closely to existing express bus routes. Figure 6(a) shows that zones in the southwestern part of the study area had a high CAI to Sadang, i.e., 0.92. They were connected to Seoul by express bus lines to Sadang. Many express lines used highways, so they were regarded as the coverage of Sadang. Figure 6(b) shows that the Gangnam transit center had a broad coverage in the southern part of the study area because Gangnam was the nearest transit center to the Gyungbu Expressway, a central corridor in the Seoul metropolitan area.

Accordingly, CAI can identify transit centers with too broad coverage that can cause overcrowding of travelers and vehicles in the center. For example, Gangnam of Seoul had extensive coverage in the southern part of Seoul, as seen in Figure 6(b). Contrarily, Jamsil did not adequately cover the southeastern part of the area in Figure 6(c). In another view, the coverage area of Sadang was too dependent on one transit center, as seen in Figure 6(a). If there are severe incidents on the road to Sadang, the areas would be isolated. The government should have other options to connect the area to Seoul with other routes or modes. This observation can suggest a solution for overcrowding in Gangnam, such as changing the destinations of regional buses from the

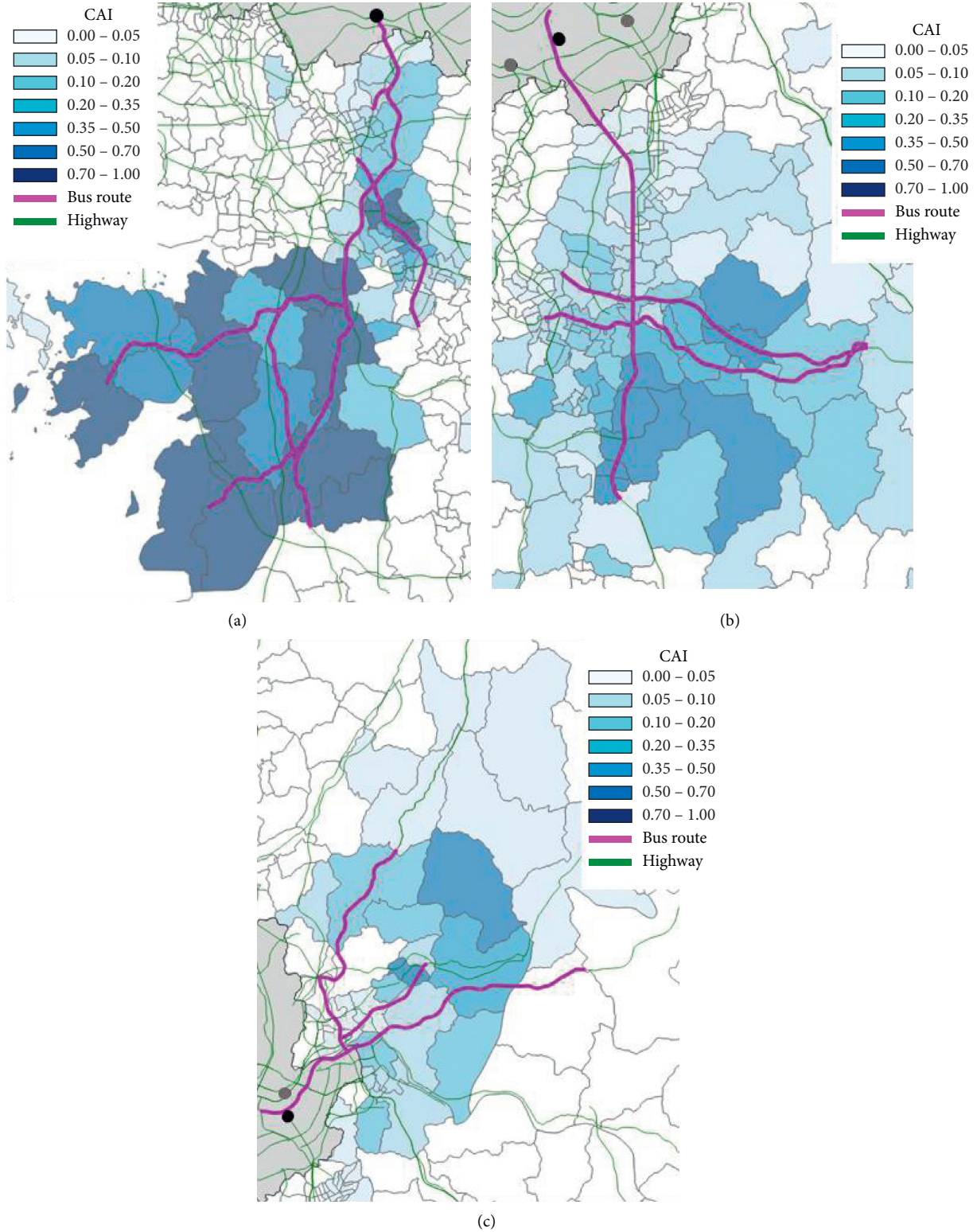


FIGURE 6: A result of CAI with regional transportation (bus routes and highway): (a) Sadang; (b) Gangnam; (c) Jamsil.

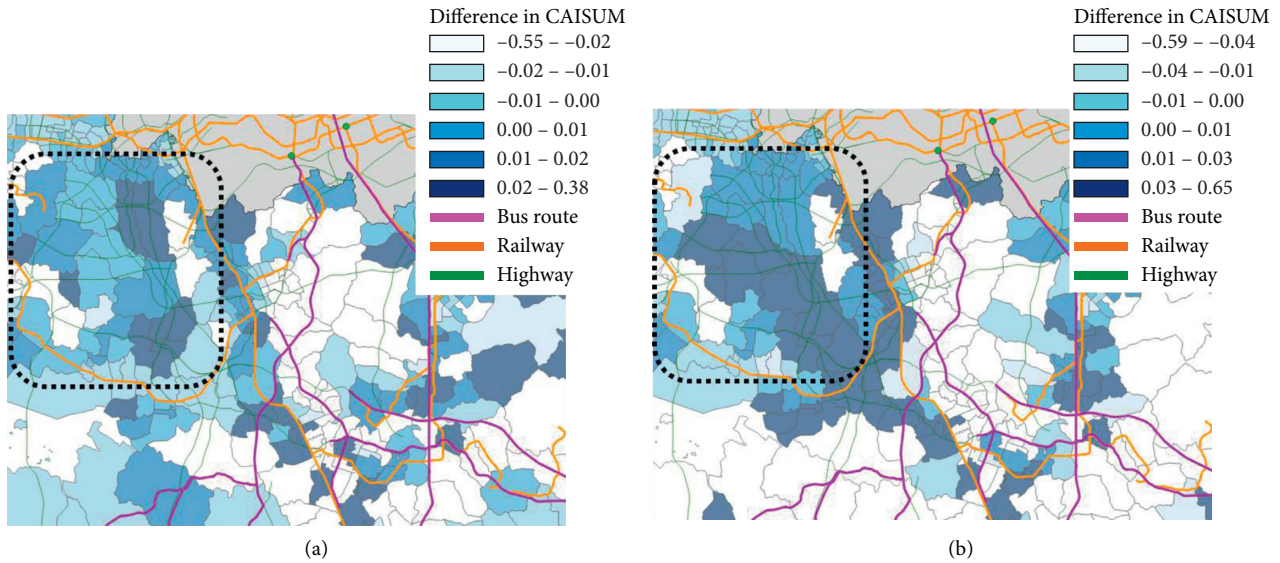


FIGURE 7: A difference in the CAISUMs between neighboring zones with regional transits (express bus routes and railway): (a) compared with zones in 5 km; (b) compared with zones in 10 km.

zones to Jamsil. With a proper supply of express bus lines, trips from the zones to the central city can be dispersed.

4.2. Connectivity between the Centers and Suburban Area. The CAIs also can be used to evaluate the regional transit in suburban zones by obtaining the sum of the CAIs for all six centers (CAISUMs). If the CAISUMs of a zone are small, it indicates that the zone is connected to the downtown through another transit center or that there is no advanced regional transit connection to the transit centers, meaning that the transit of the zone does not form hub-and-spoke structure to downtown. To distinguish between the two cases, the CAISUMs of a zone should be compared with the CAISUMs of the neighboring zones. When a zone and its neighboring zones together have small CAISUMs, it can be inferred that all of them are connected to other centers. So, comparing the CAISUM of a zone with those of its neighboring zones can be used to identify where the structure of the transit network is inefficient.

Figure 7 shows differences in the CAISUM of a zone and the CAISUMs of the neighboring zones. We used various bandwidths of KDFs to compare the CAISUMs. Differences in the CAISUM with bandwidths of 2, 5, and 10 km can show the difference in the CAISUM compared to adjusting the zones with a number. If a zone has a positive value of the difference, it had a smaller CAISUM than neighboring zones. When the railroad passes zones, the CAISUMs of the zones were calculated to be a small value. The zones were observed as having a lack of regional transit because trips from the zones to the central city by railway did not require transferring at the transit center, and the transfer between metro lines is not counted in the smart card data. The contribution of this analysis can be derived from the dotted boxes in Figure 7. These zones had lower CAISUMs than the

neighboring zones because they did not have any railways or bus routes connecting to the transit centers. Trips between the zones and the city were on city transit, which had a low level of service and not forming hub-and-spoke structure through transit centers. The contrast between Figures 7(a) and 7(b) was the size of the view. Figure 7(a) was suitable for finding the zone with an inefficient transit network in the individual zone level, and Figure 7(b) was suitable for finding the zones at the county level.

With this approach, TAZs with an insufficient service of regional transit would be identified. Travelers from the zones would experience low speed and a poor environment for transfer when they travel to downtown. The government should supply a regional express bus from the zones to existing transit centers. If the area's transit demand is too big to handle in the centers, opening a new transit center to connect subway and city buses with regional transit should be considered.

5. Conclusion

In this research, the concept of coverage was expanded to evaluate the integration and coordination of multimodal public transit system. The coverage area index (CAI) was defined as the ratio of the number of trips through the transit center to the total number of trips from the zones to the central city. Smart card data were used to conduct the case study, and the kernel density function was applied to aggregate the point-to-point origin-destination data into zones. Six transit centers that connect Seoul and the southern part of Seoul metropolitan area were selected to measure the CAIs. The peak value of CAI varied between the centers, from 0.173 for Seoul Station up to 0.92 for Sadang. CAI can be used to compare the importance of transit centers, to measure transit supply and travelers' preference and evaluate whether transit from a suburban zone to the

central city was integrated into intercity transit network of the city.

Transit centers have the role of integrating transit modes between downtown and the suburban area in the complicated hub-and-spoke transit network in big cities. Some researchers have considered the effectiveness, connectivity, and accessibility of the transit centers, but they have rarely dealt with the integration or coordination of the overall transit network. We developed CAI using the concept of betweenness centrality to evaluate the role of the centers. With the CAI, a dependence of suburban areas on the centers could be visualized. By comparing CAI and TA, we were able to evaluate CAI having the independence of the distance from the city center. Since only OD trip data were needed to measure CAI, it was not necessary to directly compare the various alternative routes and modes. The characteristics of CAI had advantages when transit networks are too complicated to analyze in the metropolitan area.

CAI could be used for three purposes: finding coverage of transit centers, measuring transit supply, and measuring connections between the centers and suburban areas. With the coverage of each center, we could find incompetent centers that had narrow and vague coverage. It meant that they could not function well as a hub, so the centers should be more attractive to travelers. With the measured supply of regional transit, alternative routes to solve overcrowding problems of centers were implicated. With the measured connection, areas with insufficient supplies of regional transit were identified, and strategies to improve the connection in the areas were suggested.

Given our conclusions and observations, future research should take into account the rarity of Seoul's smart card dataset, which provided the trip chain data we used that included the exact boarding and alighting stations for trips. The definitions suggested for coverage and CAI can be applied to other types of trip data, such as the OD matrix and transit ridership data. However, these characteristics could be limitations of the index because the index could not be used to test scenarios in the planning level due to the absence of future trip data. It can be further work to use the index with predicted future trip data. By expanding the index with the data, potential coverage areas that travelers can be users of a transit center can be identified. This result would be helpful for planners of a transit center and a transit network. In future research, analysis of coverage of transit centers could be assigned to different metropolitan areas to identify zones that have inadequate regional transit. It can be further work to use the index with the data of predicted future trips.

Data Availability

The smart card data used to support the findings of this study are restricted by the Ministry of Land, Infrastructure, and Transport of Korea in order to protect the privacy. The data are available from TRLab in Seoul National University for researchers who meet the criteria for access to confidential data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2020R1F1A1061802).

References

- [1] L. Sun, J. Rong, and L. Yao, "Measuring transfer efficiency of urban public transportation terminals by data envelopment analysis," *Journal of Urban Planning and Development*, vol. 136, no. 4, pp. 314–319, 2010.
- [2] H. Nishiuchi, T. Todoroki, and Y. Kishi, "A fundamental study on evaluation of public transport transfer nodes by data envelop analysis approach using smart card data," *Transportation Research Procedia*, vol. 6, pp. 391–401, 2015.
- [3] S. Mishra, T. F. Welch, and M. K. Jha, "Performance indicators for public transit connectivity in multi-modal transportation networks," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 7, pp. 1066–1085, 2012.
- [4] A. Ceder, Y. Net, and C. Coriat, "Measuring public transport connectivity performance applied in Auckland, New Zealand," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2111, no. 2111, pp. 139–147, 2009.
- [5] Z. Wang and Y. Chen, "Development of location method for urban public transit networks based on hub-and-spoke network structure," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2276, no. 1, pp. 17–25, 2012.
- [6] B. Yu, H. Zhu, W. Cai, N. Ma, Q. Kuang, and B. Yao, "Two-phase optimization approach to transit hub location - the case of Dalian," *Journal of Transport Geography*, vol. 33, pp. 62–71, 2013.
- [7] M.-C. Shish, H. Mahamassani, and M. Baaj, "Planning and design model for transit route networks with coordinated operations," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1623, no. 1, pp. 16–23, 1998.
- [8] E. H. Lee, H. Lee, S. Y. Kho, and D. K. Kim, "Evaluation of transfer efficiency between bus and subway based on data envelopment analysis using smart card data," *KSCE Journal of Civil Engineering*, vol. 23, no. 9, pp. 1–12, 2019.
- [9] J. G. Koenig, "Indicators of urban accessibility: theory and application," *Transportation*, vol. 9, no. 2, pp. 145–172, 1980.
- [10] S. K. Sathisan and N. Srinivasan, "Evaluation of accessibility of urban transportation networks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1617, no. 1, pp. 78–83, 1998.
- [11] Kittelson & Associates, USFT Administration, TCR Program, and TD Corporation, *Transit Capacity and Quality of Service Manual*, Transportation Research Board, Washington, DC, USA, 2003.
- [12] S. A. Mamun and N. E. Lownes, "Measuring service gaps: accessibility-based transit need index," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2217, no. 1, pp. 153–161, 2011.
- [13] I. Tasic, C. Bozic, and E. Hanss, "Towards true multimodal transportation accessibility: data, measures, and methods," in *Proceedings of the 96th Annual Meeting of the Transportation Research Board (TRB)*, Washington, DC, USA, January 2017.

- [14] C.-L. Cheng and A. W. Agrawal, "TTSAT: a new approach to mapping transit accessibility," *Journal of Public Transportation*, vol. 13, no. 1, pp. 55–72, 2010.
- [15] B. M. Alam, G. L. Thompson, and J. R. Brown, "Estimating transit accessibility with an alternative method: evidence from Broward County, Florida," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2144, no. 1, pp. 62–71, 2010.
- [16] K. Sohn and D. Kim, "Zonal centrality measures and the neighborhood effect," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 9, pp. 733–743, 2010.
- [17] S. B. Chung, S. H. Hur, J. H. Baek, D. Kim, and K. H. Song, "Development of accessibility index using closeness centrality," in *Proceedings of the Eastern Asia Society for Transportation Studies the 9th International Conference of Eastern Asia Society for Transportation Studies*, Eastern Asia Society for Transportation Studies, Jeju, Korea, 2011.
- [18] X. Ma and Y. Wang, "Development of a data-driven platform for transit performance measures using smart card and GPS data," *Journal of Transportation Engineering*, vol. 140, no. 12, 2014.
- [19] E.-J. Kim, Y. Kim, and D.-K. Kim, "Interpretable machine-learning models for estimating trip purpose in smart card data," in *Proceedings of the Institution of Civil Engineers—Municipal Engineer*, Thomas Telford Ltd, pp. 1–22, 2020.
- [20] W. Tu, R. Cao, Y. Yue, B. Zhou, Q. Li, and Q. Li, "Spatial variations in urban public ridership derived from GPS trajectories and smart card data," *Journal of Transport Geography*, vol. 69, pp. 45–57, 2018.
- [21] H. Lee, H. Park, S. Kho, and D. Kim, "Assessing transit competitiveness in Seoul considering actual transit times based on smart card data," *Journal of Transport Geography*, vol. 80, Article ID 102546, 2019.
- [22] C. Kapuku, S.-Y. Kho, D.-K. Kim, and S.-H. Cho, "Modeling the competitiveness of a bike-sharing system using bicycle GPS and transit smartcard data," *Transportation Letters*, pp. 1–5, 2020.
- [23] Z. Cai, D. Wang, and X. M. Chen, "A novel trip coverage index for transit accessibility assessment using mobile phone data," *Journal of Advanced Transportation*, vol. 2017, Article ID 9754508, 14 pages, 2017.
- [24] M. T. Magalhães, "Spatial coverage index for assessing national and regional transportation infrastructures," *Journal of Transport Geography*, vol. 56, pp. 53–61, 2016.
- [25] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
- [26] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1978.
- [27] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical Review Letters*, vol. 87, 2001.
- [28] G. L. Thompson, "Achieving suburban transit potential: sacramento revisited," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1571, no. 1, pp. 151–160, 1997.
- [29] Korea Transport DataBase, *Establishing Network for Transport Analysis 2018*, The Korea Transport Institute and Ministry of Land, Infrastructure and Transport of Korea, Sejong City, South Korea, 2018.

Research Article

Person Detection for an Orthogonally Placed Monocular Camera

Pavel Skrabanek ¹, Petr Dolezel,² Zdenek Nemecek,² and Dominik Stursa²

¹Faculty of Mechanical Engineering, Brno University of Technology, Brno, Czech Republic

²Faculty of Electrical Engineering and Informatics, University of Pardubice, Pardubice, Czech Republic

Correspondence should be addressed to Pavel Skrabanek; pavel.skrabanek@vut.cz

Received 6 May 2020; Revised 15 July 2020; Accepted 23 September 2020; Published 14 October 2020

Academic Editor: Kun Xie

Copyright © 2020 Pavel Skrabanek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Counting of passengers entering and exiting means of transport is one of the basic functionalities of passenger flow monitoring systems. Exact numbers of passengers are important in areas such as public transport surveillance, passenger flow prediction, transport planning, and transport vehicle load monitoring. To allow mass utilization of passenger flow monitoring systems, their cost must be low. As the overall price is mainly given by prices of the used sensor and processing unit, we propose the utilization of a visible spectrum camera and data processing algorithms of low time complexity to ensure a low price of the final product. To guarantee the anonymity of passengers, we suggest orthogonal scanning of a scene. As the precision of the counting is relevantly influenced by the precision of passenger recognition, we focus on the development of an appropriate recognition method. We present two opposite approaches which can be used for the passenger recognition in means of transport with and without entrance steps, or with split level flooring. The first approach is the utilization of an appropriate convolutional neural network (ConvNet), which is currently the prevailing approach in computer vision. The second approach is the utilization of histograms of oriented gradients (HOG) features in combination with a support vector machine classifier. This approach is a representative of classical methods. We study both approaches in terms of practical applications, where real-time processing of data is one of the basic assumptions. Specifically, we examine classification performance and time complexity of the approaches for various topologies and settings, respectively. For this purpose, we form and make publicly available a large-scale, class-balanced dataset of labelled RGB images. We demonstrate that, compared to ConvNets, the HOG-based passenger recognition is more suitable for practical applications. For an appropriate setting, it defeats the ConvNets in terms of time complexity while keeping excellent classification performance. To allow verification of theoretical findings, we construct an engineering prototype of the system.

1. Introduction

In passenger transport, person flow monitoring has an indispensable importance. In some areas of public transport, passenger flow monitoring systems are employed to automate this task. One of the basic measures, which must be provided by the system, is the number of transported passengers. A precise counting of passengers entering and exiting means of transport has a positive effect on public transport surveillance, passenger flow prediction, transport planning, transport vehicle load monitoring, station control and management, and cost optimization [1, 2].

To ensure a robust and precise counting of passengers in real time, a passenger flow monitoring system must be based

on an appropriate imaging system and data processing algorithms. In order to allow a mass deployment of such a monitoring system, a low-cost final solution is equally important. The solution should also meet legal requirements where passenger anonymity is of great importance. Specifically, identification of individuals according to their faces must be avoided.

The imaging system must ensure the acquisition and processing of data, i.e., its basic components are a sensor and a processing unit. In order to develop an inexpensive solution, low price of both components is crucial. While the lower price limit of the processing unit is mainly given by the complexities of used data processing algorithms, the lower price limit of the sensor is given by the used sensing

technology. Radar sensors [3], laser scanners [4], 3D laser scanners [5], or infrared sensors [6] are applicable for the counting of passengers. All these sensors naturally guarantee a high level of passenger anonymity. Their main drawbacks are high prices of the sensors and frequent errors in the counting [7, 8]. For these reasons, cameras operating in the visible spectrum of light are preferably used for the counting of persons [9]. Conventional cameras (cameras operating at wavelengths of visible light) are significantly cheaper, compared to the previously mentioned sensors. The cameras can be combined with depth sensing devices [10]. The fusion of data can result in a more balanced trade-off between false positives and false negatives [11]. On the other hand, the depth sensing devices increase the final prices of sensors, i.e., utilization of a depth sensing device would increase the final price of the imaging system.

The automated counting of persons in a scene is usually carried out in colour images or in sequences of colour images. Many data processing algorithms aimed at precise counting of persons in crowded scene images have been presented [9]. Most of them are designed for overriding installations of cameras. Cameras installed at public as well as at private places usually look down on scenes from angles that typically range between 40° and 80° (from the ground). Considering low subject distances in transportation means (a distance between a camera and a passenger), we conclude that the anonymity of passengers is not guaranteed for such a setup (i.e., data processing algorithms aimed at processing of such images cannot be used for the counting of passengers). Only orthogonally captured images (camera placed above a scene, looking directly down on the scene) assure a high level of passenger anonymity (Figure 1).

A data processing chain, aimed at counting of persons in orthogonally captured images, is compounded of three fundamental steps: person detection, multiperson tracking, and person counting (Figure 2). In the first step, a processed image is examined for the presence of persons. The following step is the tracking, where all persons detected within the first step are matched with existing tracking models of persons. In the case a person cannot be associated with any existing model, a new tracking model is initialized. The last step of the chain is the counting. If a person described by a tracking model leaves the scene, which is usually defined by virtual lines, counting is triggered [11]. Naturally, an integral part of this data processing chain is an algorithm which splits video data provided by a camera into individual images.

Accuracy and time complexity of the data processing chain is primarily given by accuracy and time complexity of the person detection. Person detection is a process of location and recognition of persons in images. Within this process, possible locations of persons (regions) are proposed using an appropriate technique. The regions determine candidate object images, which are classified using an appropriate object recognition system. The proposition of regions can be carried out using an exhaustive method such as a sliding window [12] or using an advanced time-efficient method such as a selective search algorithm [13]. In modern object detection systems, both the location and the

recognition are carried out by a single deep neural network [14–16]. These systems are characterized by high detection accuracy but high time complexity.

As the analysis shows, a low-cost passenger counting system should be based on a conventional camera (due to low prices of visible light cameras). In order to guarantee the anonymity of passengers, the camera must be placed above a scene, looking directly down on the scene. For the data processing, methods capable of processing orthogonally captured images must be used. The resulting data processing chain must be robust and precise. To keep the low-cost requirement, the time complexity of the methods should be as low as possible. From this perspective, the detection of passengers seems to be the weak link in the chain.

As the time complexity of the single deep neural network detectors is high [14–16], we tend to implement a passenger detector as a two-stage system. When using a robust and time-efficient region proposal method such as selective search algorithm [13], the accuracy and computational complexity of the detector is mainly given by a used object recognition method. In colour images, the recognition of persons typically relies on optical flow features [11, 17, 18]. An alternative approach to the detection of persons is the detection of their heads and shoulders [19]; however, a head itself can provide a strong feature due to its almost circular shape. The counting of heads is typically used by counting of persons in dense crowd images [20, 21]. Recognition of heads in orthogonally captured images can also rely on the optic flow analysis [22]. The main disadvantages of optic flow-based methods are their high computational complexity and noise sensitivity [23].

Considering the importance of passenger recognition for their counting, we focus on the development of a price-competitive and time-efficient object recognition system. As the system is aimed at recognition of passengers, we name it “the passenger recognition system.” As the trend in object recognition is still clearly heading towards convolutional neural networks (ConvNets) [24, 25], we examine the performance of ConvNets for passenger recognition. Usually, ConvNet-based object recognition systems have good classification performance, but their time complexity is typically high. For this reason, we propose a competitive approach which is based on histograms of oriented gradients (HOG) features [26] and on a support vector machine (SVM) classifier. For an appropriate setting of parameters, HOG-based object recognition can have good classification performance while keeping low time complexity [27].

Recognition of passengers in orthogonally captured images using the HOG features and the SVM classifier, based on object images which comprise of heads and shoulders of passengers, has proven useful in scenes without height differences [19]. Modern public means of transport are increasingly low-floor (i.e., there is no or negligible height difference in the area of a doorway), but a substantial part of operated buses, trams, trains, and trolleybuses are high-floor [28–30]. Considering this fact, we conclude that the robustness of the HOG-based passenger recognition system must be verified in the context of variable distances between a camera lens and passenger heads. We also consider that the



FIGURE 1: Examples of images orthogonally captured in a tram. Identification of persons in the images according to their faces is implausible.

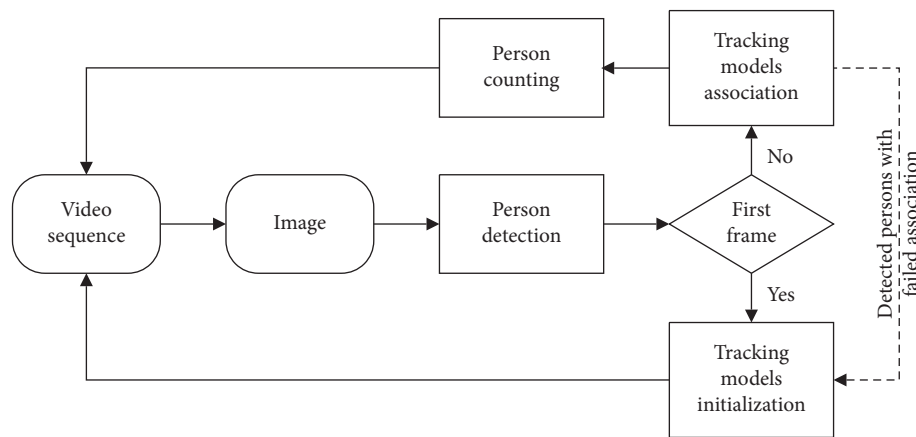


FIGURE 2: Overall diagram of a person counting system. The system processes images streamed by a monocular camera. Each image is firstly examined for a presence of persons. Positive samples are then associated with existing tracking models of persons. If sample is not associated with any existing model, a new model is initialized. If a tracking model leaves scene, which is defined by virtual lines, counting is triggered.

time complexity of the system can be reduced once the object image contains only the heads of the passengers (omitting the shoulders will result in smaller object images and consequently reduce data processing time). We deduce the suitability of such an approach from remarkable results of HOG-based object recognition systems on similar tasks, e.g., for grape detection [31, 32] (see Figure 3; the round shape of grapes is similar to the shape of heads).

Within this article, we study the classification performances and time complexities of passenger recognition systems. The systems are aimed at recognition of passengers in orthogonally captured images, where the recognition quality is not adversely affected by the variable distance between the passenger and camera sensor. The passenger recognition systems are based either on ConvNets or on HOG features. Both approaches rely on the detection of heads. In the case of ConvNet-based systems, we consider ConvNets of various topologies. In the case of the HOG-based system, we examine various settings of parameters. We validate the theoretical results in a real-world application. For this purpose, we develop an engineering prototype of the system.

2. Material and Methods

2.1. Engineering Prototype of the System. Two basic components of the system are the sensor and the processing unit (Figure 4). In our case, we use an industrial colour camera Basler acA2500-60uc [34] as the sensor. The camera is placed in a means of transport, at the ceiling near a door. The optical axis of the camera is perpendicular to the vehicle floor. Considering the construction of means of transport, we expect the average subject distances to be from 0.2 m to 1 m. The camera should monitor an area of about 2.4 m × 2.0 m. With respect to these parameters, we equipped the camera with a Computar M3514-MP lens [35]. The output of the camera (i.e., the input of the data processing chain) is a sequence of RGB images.

We use the prototype for a data collection as well as for the validation of the proposed recognition methods, i.e., the prototype must be capable of processing acquired images in real time. In order to allow testing of all proposed solutions (including solutions based on ConvNets), we use a single-board computer VOB-P3310. It offers an NVIDIA Tegra X2 (2.0 GHz, 6 cores) CPU together with 8 GB RAM and it

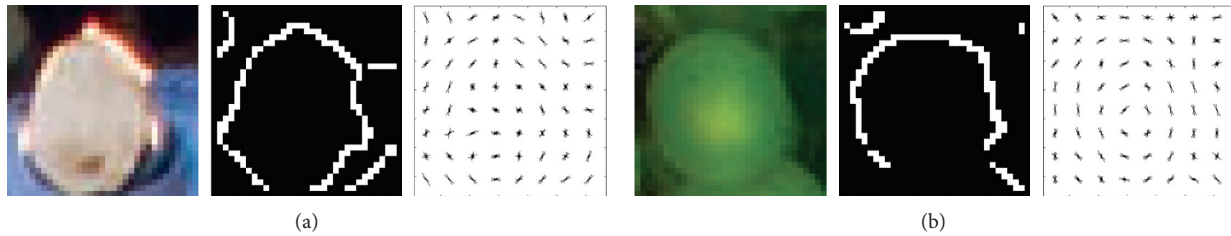


FIGURE 3: Comparison of head (three images in (a)) and grape images (three images in (b)). For each category, we provide an original RGB image, an image obtained by filtering of the RGB image using the Canny edge detector [33], and gradients obtained using a HOG descriptor [26], respectively.

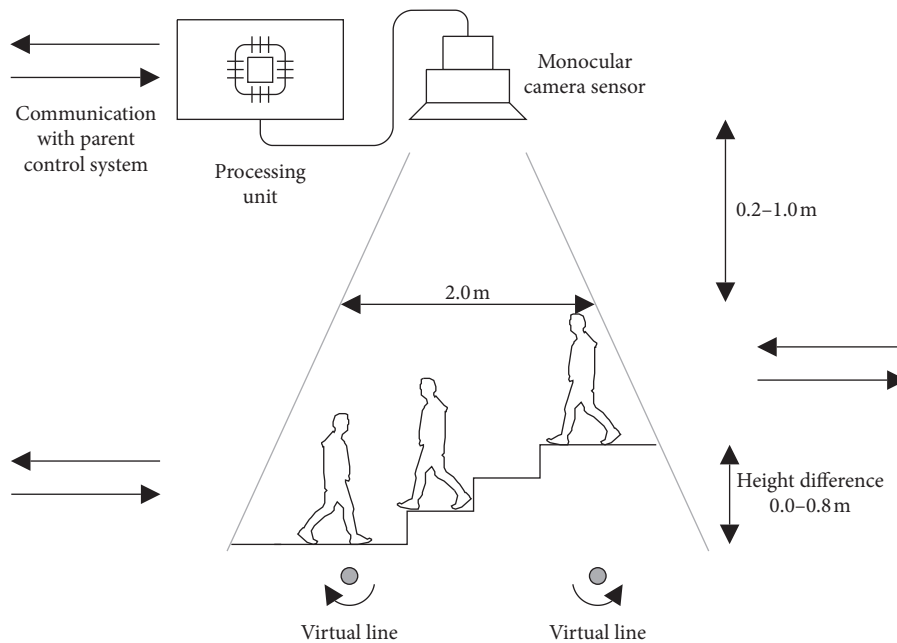


FIGURE 4: Architecture of person flow monitoring system (side view). Crossing space between lines causes person to be counted.

provides wide communication possibilities (USB 2.0, 3.0, SATA, WiFi) [36].

2.2. Passenger Recognition. Candidate object images may or may not contain complete heads of passengers (Figure 5). According to this criterion, the images are classified either as “head” or “not head” by a passenger recognition system. Inputs of the recognition system are sized normalized RGB object images of dimensions 51×51 pixels ([51, 51] px). Its outputs are labels of the images, where labels “head” and “not head” are allowed.

2.2.1. Passenger Recognition Based on ConvNets. In terms of classification accuracy, the state-of-the-art object recognition systems are based on one of the successful deep ConvNet architectures [37]. Mostly, they process raw image data (i.e., no image preprocessing is usually carried out). They consist of multiple layers arranged in a feed-forward manner. Upper and lower level layers ensure feature extraction and classification of object images, respectively. The

feature extraction is usually carried out using convolutional and pooling layers, where the convolutional layers are typically combined with a ReLU activation function. The classification is generally ensured by a softmax activation function. The function processes data at the output of the last network layer, where a fully connected layer is placed. The number of neurons of this layer corresponds to the number of object classes [38]. The main drawback of the state-of-the-art deep ConvNet architectures is their high computational demands.

The passenger recognition can be simply implemented as a ConvNet of an appropriate architecture, where the network ensures both feature extraction and classification (Figure 6). As a low time complexity of the system is crucial, we test the performance of five ConvNet architectures of different complexities.

The simplest architecture, Net1, consists of one convolutional layer (32 filters with 3×3 px kernels), one max-pooling layer (2×2 px nonoverlapping pools), and two fully connected layers of 512 and 2 neurons, respectively. The classification is carried out using the softmax function. In the

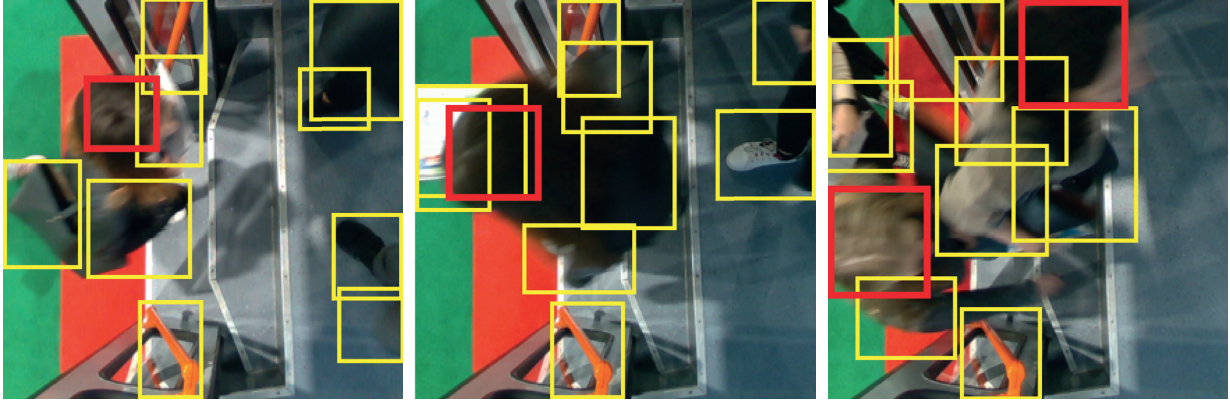


FIGURE 5: Detection of passengers in orthogonally acquired video data. Candidate object images proposed by a search algorithm (yellow rectangles) are tested. Some of them are classified as “head” (red rectangles), which represents the locations of passengers.

second simplest architecture, Net2, we replace the convolutional and the max-pooling layers by the sequence of layers: convolutional layer ($32, 3 \times 3$) + convolutional layer ($32, 3 \times 3$) + max-pooling layer + convolutional layer ($64, 3 \times 3$) + convolutional layer ($64, 3 \times 3$) + max-pooling layer, where 2×2 px nonoverlapping pools are used at both pooling layers. In both networks, we use ReLU activation functions at the convolution and fully connected layers. To reduce overfitting, we place dropout layers after each max-pooling layer and after the first fully connected layers in both networks. The dropout rate is 25% and 50% for the max-pooling and the fully connected layers, respectively.

The remaining three architectures studied within this article are the well-known LeNet-5 [39, 40], AlexNet [41], and VGG-16 net [42]. The networks are ordered according to their complexities. The LeNet-5 is the pioneering ConvNet of a relatively simple architecture. AlexNet is probably the most cited deep ConvNet with a huge number of industrial and engineering applications. VGG-16 is a representative of very deep ConvNets. As it consists of only 13 and 3 convolutional and fully connected layers, respectively, the real-time processing of data by VGG-16 implemented in the engineering prototype (Section 2.1) is still possible.

We train all the networks from scratch with initial weights set randomly with normal distribution (mean = 0, standard deviation = 0.05). In addition, we use transfer learning (TL) for AlexNet and VGG-16 in order to test the possibility of better performance [41, 42]. For both architectures, we fine-tune the last three layers of the pretrained networks.

Due to a stochastic character of the training process, we repeat the training a hundred times for each network and training strategy. For each training, we randomly split up a training set into training and validation subsets at the ratio 85 : 15. For each training subset, we run the training in a batch mode for 100 epochs with batches of 32 images. We randomly shuffle data in training subsets for each epoch. We use an ADAM optimizer [43] with initial learning rate setup at 10^{-3} and exponential decay rates for the first and second moment estimates setup at 0.9 and 0.999, respectively. The optimizer and setting of the hyperparameters are the results of a pilot study. We minimize a binary cross-entropy function:

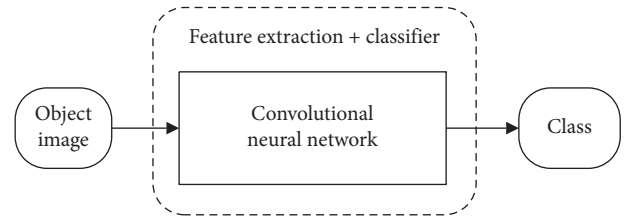


FIGURE 6: Vision pipeline of the passenger recognition systems based on ConvNets.

$$E_{\text{ConvNet}} = -\frac{1}{n} \sum_{j=1}^n [\hat{y}_j \ln(y_j) + (1 - \hat{y}_j) \ln(1 - \hat{y}_j)], \quad (1)$$

where n is the number of images in the training subset and y_j and \hat{y}_j are an actual and a predicted class of the j -th object image, respectively. We validate each such trained network on the corresponding validation subset using the cross-entropy function (1).

2.2.2. Passenger Recognition Based on HOG and SVM.

Herein, we present a passenger recognition system developed using traditional computer vision techniques. A vision pipeline of the system consists of three successive steps: image preprocessing, feature extraction, and classification (Figure 7). For the feature extraction and classification, we use the HOG descriptor and SVM classifier, respectively. In order to reduce the time complexity of the system, we convert input RGB images to the grayscale format within the image preprocessing. The conversion is carried out according to the ITU-R recommendation BT.601 [44]. The second step of the preprocessing is the unity-based normalization of the grayscale images [31].

The HOG descriptor encodes local shape information from regions within an image into a feature vector [26]. The descriptor has five parameters: number of bins, orientation binning, size of cells (in pixels), number of cells in blocks, and number of overlapping cells between adjacent blocks. As the size of cells significantly influences the final performance of image recognition systems [27] (Figure 8), we study the

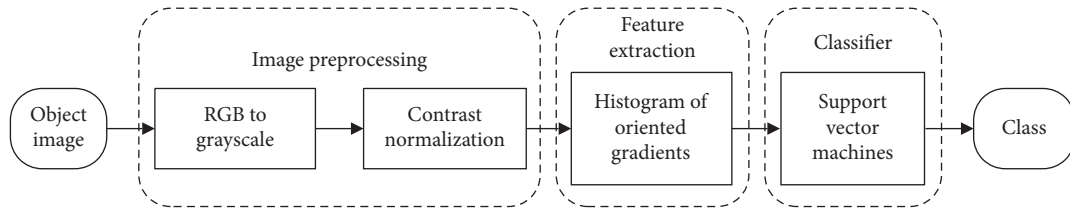


FIGURE 7: Vision pipeline of the passenger recognition system based on the HOG descriptor and on the SVM classifier. The conversion combined with the normalization is aimed at reducing of time complexity of the system.



FIGURE 8: From left to right: the original object image [51, 51] px and images with highlighted gradients of HOG features for cell of sizes [16, 16] px, [8, 8] px, and [6, 6] px, respectively. The length of white abscissae is related to the gradients in the image.

influence of this parameter on the classification performance of the HOG-based passenger recognition system. Specifically, we consider cells of sizes [6, 6], [8, 8], ..., [16, 16] px. For the remaining parameters, we use a conservative setting which has proven to be efficient: linear gradient voting into 9 bins linearly spread over 0 to 180 degrees, blocks of 2×2 cells, and 1 overlapping cell between adjacent blocks in both directions.

Training of the SVM classifier is an optimization problem which searches for a hyperplane with a maximal margin from the training data [45]. In the case that the data is not linearly separable, the data must be transformed into a linearly separable problem using an appropriate kernel function. For strongly nonlinear problems, selection of the kernel function is crucial. Considering this fact, we test the influence of various kernels on the performance of the HOG-based passenger recognition system. Specifically, we focus on the well-established linear, Gaussian radial basis function (RBF), and polynomial kernel functions (we use polynomial kernel with order equal to 2 and 3).

Performances of SVM classifiers are also influenced by settings of their regularization constants. In the case that an SVM classifier uses the RBF kernel, its performance is further influenced by kernel width. In a pilot study, we have found setting of the regularization constant at 1 to be optimal. We use a subsampling-based heuristic procedure to find the optimal setting of the kernel width.

As classification performances of classifiers strongly depend on the composition of training sets, we search for a setting ensuring the best performance of the HOG-based passenger recognition system. We carry out the search in the manner described in Section 2.2.1. Specifically, we randomly split up the training set into training and validation subsets at the ratio 85 : 15, and we train and validate the system on

the subsets. We repeat the training-validation process a hundred times for each possible combination of kernel function and cell size. We carry out the validation on corresponding validation subsets using a loss function that is given as a sum of misclassified observations, i.e.,

$$E_{\text{SVM}} = \sum_{j=1}^n I\{\hat{y}_j \neq y_j\}, \quad (2)$$

where $I\{\cdot\}$ is the indicator function.

2.3. Evaluation of Passenger Recognition. Two key aspects of the presented passenger recognition systems are their classification performances and their time complexities. A common practice of the evaluation of the classification performance is calculation of accuracy over a testing set (a dataset independent of the training set). For the classification of images into categories “positive” and “negative,” the accuracy is given as follows:

$$\text{accuracy} = \frac{|\text{TP}| + |\text{TN}|}{|\text{TP}| + |\text{FP}| + |\text{TN}| + |\text{FN}|}, \quad (3)$$

where |TP| is the number of correctly classified positive images, |FN| is the number of misclassified positive images, |FP| is the number of misclassified negative images, and |TN| is the number of correctly classified negative images.

To evaluate the classification performance comprehensively, we use three additional measures [31, 46]:

$$\text{precision} = \frac{|\text{TP}|}{|\text{TP}| + |\text{FP}|}, \quad (4)$$

$$\text{recall} = \frac{|\text{TP}|}{|\text{TP}| + |\text{FN}|}, \quad (5)$$

$$F1 - \text{score} = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}}. \quad (6)$$

To evaluate the time complexities of the systems, we measure times that the systems needed to process the testing set. To keep the results independent on the used hardware, we operate with a relative computational time. For the j -th evaluated system, its relative computational time is given as follows:

$$T_j = \frac{t_j}{\max\{t_1, t_2, \dots, t_k\}} \times 100, \quad (7)$$

where t_j is time the j -th system needs to process the data and k is the number of all evaluated systems.

We carry out the evaluation of passenger recognition systems using the best models obtained within the training process (see Sections 2.2.1 and 2.2.2). In the case of ConvNet-based systems, we use for each architecture, the model with the smallest value of the cost function (1) obtained by its validation. In the case of the HOG-based system, we use for each setting the model with the smallest value of the cost function (2) obtained by its validation.

2.4. Training and Testing Sets. Quality and composition of the training and testing sets conspicuously influence the overall performance of object recognition systems in real-life applications. Data included in the sets should reflect as many aspects of the real situation as possible. Considering this fact, we base the sets on video sequences acquired in the means of public transport and similar public places under various light conditions, using the engineering prototype.

A set of candidate object images generated by a search algorithm from a frame is imbalanced (often highly) [12, 13] with a predominance of images without complete heads (Figure 5). As conventional SVMs are not suitable for the imbalanced learning tasks [47], the training and testing sets must be balanced to get unbiased results. Considering these facts, we create the sets manually to ensure the balance of the classes in the sets.

Specifically, we perform four distinct video recording experiments. They are set to simulate the real situation as well as to comprehend the architecture of the assumed person flow monitoring system (see Figure 4). All the experiments include stairs and a group of persons walking under the acquisition sensor. Men, women, and children as well as people with and without a head cover (hats, scarves, caps, and hoods) are included. Since the used camera lens is focused manually (once for each experiment), the acquired frames show certain blurring according to the specific distance between the object and the lens. We varied locations, lighting conditions, number of frames, mean distance between persons \bar{D}_{pp} (mean distance between a subject and two other nearest persons), and minimal and maximal distances between a head and the sensor, $\min D_{HS}$ and $\max D_{HS}$, in each experiment (Table 1).

We cut out and size normalize 6020 unique object images from the video data (dimension of the normalized images are [51, 51] px). We label the images according to the

presence/absence of heads (Figure 9). We mix and divide the labelled images into the training and testing sets according to Table 2. We make the sets publicly available at [48]. The sets contain large-scale class-balanced data which make them universally applicable (the sets can be used to design any classifier including classifiers, which are not suitable to be trained with imbalanced training sets).

3. Results

3.1. Validation of Passenger Recognition Systems. We train and validate each proposed architecture (ConvNet-based system) and each setting (HOG-based system) a hundred times. To show the validation results, we use box plots. Results obtained for the systems based on ConvNets are shown in Figure 10. The central lines in the graphs are medians of the loss function (1); the edges of the boxes are 25th and 75th percentiles; and the whiskers indicate the variability outside the upper and lower quartiles. The data are grouped according to the architectures and training strategies (x -axis). The values on the y -axis correspond to the loss function values.

Figure 11 shows validation results obtained for the HOG-based passenger recognition system using the loss function (2). We use a separate graph for each kernel function. Data in the graphs are grouped according to the sizes of cells. Outliers are symbolized using stars.

3.2. Classification Performance of Passenger Recognition Systems. In Table 3, we summarize evaluation results obtained from the testing set using the measures (3)–(6). The results are grouped into two sections according to the approach they are based on. The best results obtained for each measure are in bold for both approaches.

3.3. Time Complexities of Passenger Recognition Systems. We display relative computational times (7) as a bar graph (the lower chart in Figure 12), where the time and evaluated systems are on the y - and x -axes, respectively. Above each result, we display the $F1$ -score (6) of the system as a bar graph (the upper chart in Figure 12), where the $F1$ -score and evaluated systems are on the y - and x -axes, respectively.

4. Discussion

The main objective of the presented work is comparison of the two well-established object recognition approaches for the passenger recognition task. As the evaluation results (Table 3) show, for the cells of size [10, 10] px and the polynomial kernel function of degree 3, the classification performance of the HOG-based system slightly exceeds the classification performance of ConvNet-based systems. For this setting, the HOG-based system has the highest values of all four measures. The ConvNet-based systems show the best results for only one measure at a time (aside from LeNet-5 with highest accuracy and $F1$ -score). Except for recall, the HOG-based system also exceeds the ConvNet-based systems in sizes of the performance measure values. Further, for this

TABLE 1: Parameters of the video recording experiments. For each experiment, a location, lighting conditions, number of frames, mean distance between persons \bar{D}_{pp} , and minimal and maximal distances between a head and the sensor, $\min D_{HS}$ and $\max D_{HS}$, are specified. Note that individual persons can be present in multiple frames in different positions.

Location	Light	No. of frames	\bar{D}_{pp}	$\min D_{HS}$	$\max D_{HS}$
Outdoors	Ambient, strong	1720	0.00–0.50	1.2	1.6
Indoors	Ambient, strong	1700	0.25–0.75	0.6	1.8
Indoors	Ambient, weak	1400	0.25–0.75	0.4	1.4
Indoors	Artificial, weak	1200	0.50–1.00	0.2	1.0



FIGURE 9: Examples of object images in the sets. The first three images (a) are labelled as “head” while the remaining three (b) are labelled as “not head.”

TABLE 2: Dataset.

Set	Training		Testing	
Class	“Head”	“Not head”	“Head”	“Not head”
No. of images	2008	2012	1000	1000

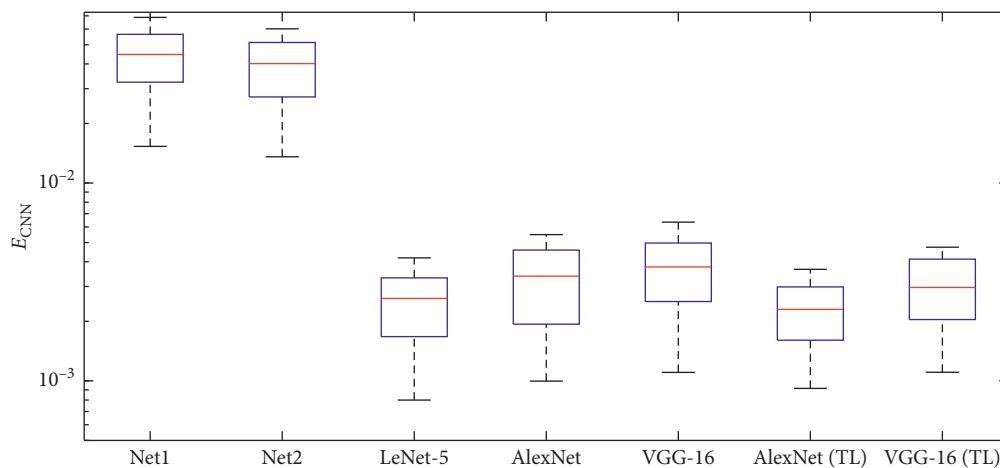


FIGURE 10: Boxplot representation of loss function values (y -axis) obtained by validation of ConvNet-based passenger recognition systems. Architectures of the ConvNets and training strategy are at x -axis (TL = transfer learning, otherwise, the network was trained from scratch).

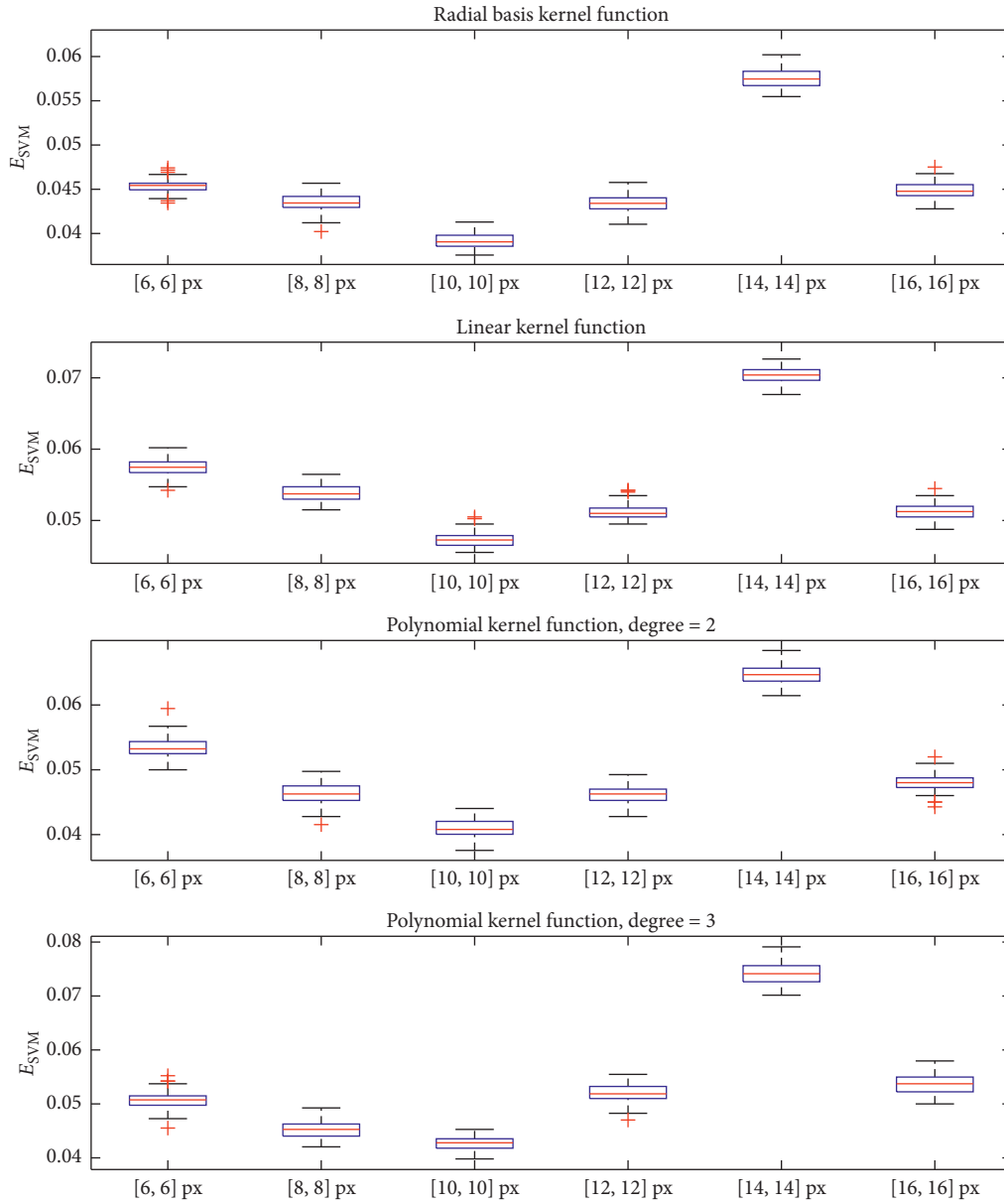


FIGURE 11: Boxplot representation of loss function values (y -axis) obtained by validation of the HOG-based passenger recognition system. Results are displayed in separate graphs with respect to used kernel function. In each graph, the data are grouped according to the sizes of cells (x -axis).

TABLE 3: Evaluation of classification performance of ConvNet-based (first section) and HOG-based (second section) passenger recognition systems using the measures (3)–(6).

Classifier	Accuracy	Precision	Recall	F1-score
Net1	0.949	0.950	0.948	0.949
Net2	0.953	0.947	0.961	0.954
LeNet-5	0.956	0.946	0.966	0.956
AlexNet	0.947	0.921	0.977	0.948
VGG_16	0.928	0.903	0.958	0.930
SVM with RBF kernel function, cell size [6, 6] px	0.949	0.957	0.941	0.949
SVM with linear kernel function, cell size [6, 6] px	0.939	0.947	0.931	0.939
SVM, polynomial degree = 2, cell size [6, 6] px	0.946	0.946	0.947	0.946
SVM, polynomial degree = 3, cell size [6, 6] px	0.953	0.953	0.947	0.950
SVM with RBF kernel function, cell size [8, 8] px	0.949	0.952	0.945	0.948
SVM with linear kernel function, cell size [8, 8] px	0.938	0.941	0.934	0.937

TABLE 3: Continued.

Classifier	Accuracy	Precision	Recall	F1-score
SVM, polynomial degree = 2, cell size [8, 8] px	0.947	0.951	0.943	0.947
SVM, polynomial degree = 3, cell size [8, 8] px	0.948	0.949	0.946	0.948
SVM with RBF kernel function, cell size [10, 10] px	0.956	0.964	0.947	0.956
SVM with linear kernel function, cell size [10, 10] px	0.943	0.946	0.939	0.943
SVM, polynomial degree = 2, cell size [10, 10] px	0.947	0.947	0.948	0.948
SVM, polynomial degree = 3, cell size [10, 10] px	0.959	0.957	0.961	0.959
SVM with RBF kernel function, cell size [12, 12] px	0.953	0.956	0.948	0.952
SVM with linear kernel function, cell size [12, 12] px	0.935	0.939	0.930	0.934
SVM, polynomial degree = 2, cell size [12, 12] px	0.950	0.955	0.945	0.950
SVM, polynomial degree = 3, cell size [12, 12] px	0.950	0.957	0.942	0.949
SVM with RBF kernel function, cell size [14, 14] px	0.929	0.925	0.913	0.919
SVM with linear kernel function, cell size [14, 14] px	0.919	0.925	0.913	0.919
SVM, polynomial degree = 2, cell size [14, 14] px	0.929	0.936	0.922	0.929
SVM, polynomial degree = 3, cell size [14, 14] px	0.921	0.920	0.923	0.922
SVM with RBF kernel function, cell size [16, 16] px	0.952	0.955	0.949	0.952
SVM with linear kernel function, cell size [16, 16] px	0.942	0.949	0.934	0.941
SVM, polynomial degree = 2, cell size [16, 16] px	0.948	0.951	0.944	0.948
SVM, polynomial degree = 3, cell size [16, 16] px	0.943	0.943	0.943	0.943

Note: best results are in bold.

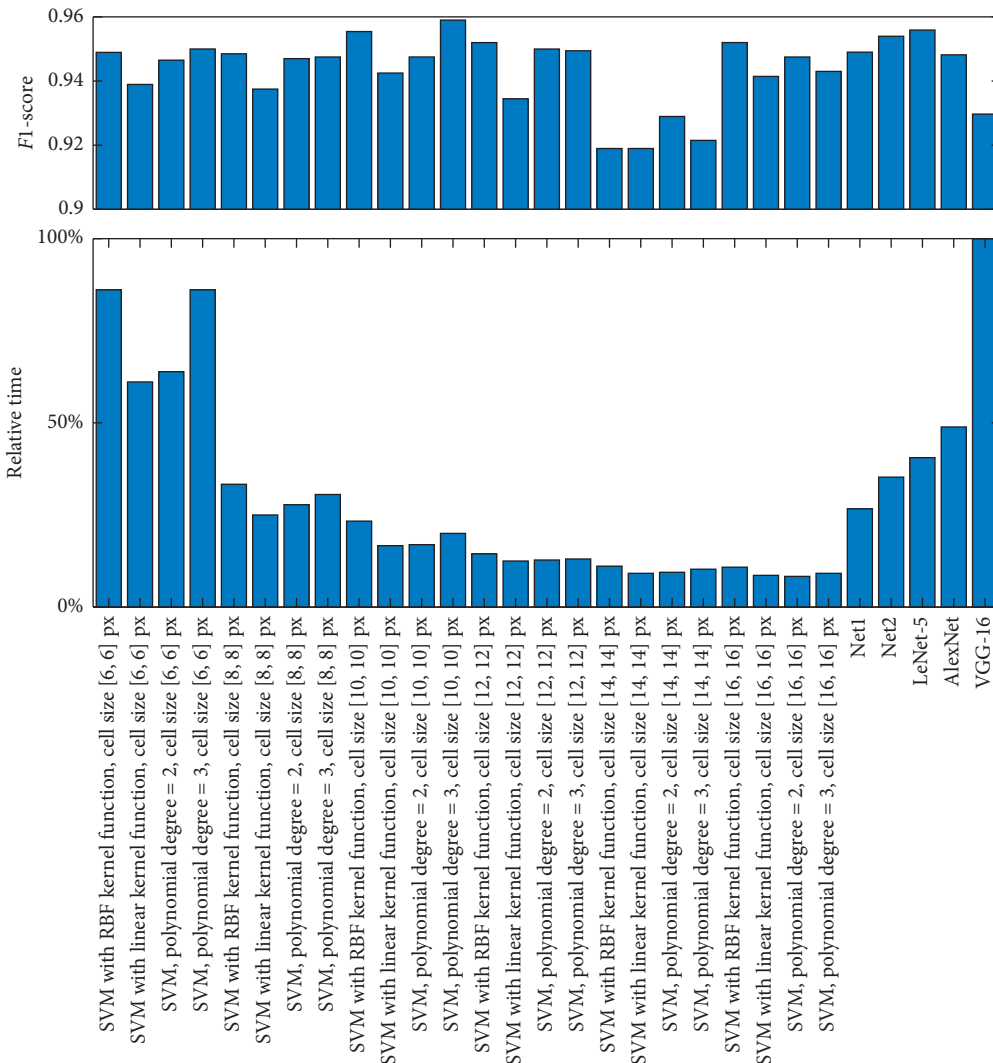


FIGURE 12: Relative computational times of the passenger recognition systems in comparison with F1-score of each system.

setting, the HOG-based system has significantly lower time complexity when compared to the ConvNet-based systems (Figure 12). Considering all these facts, we conclude that the HOG-based passenger recognition system, with polynomial kernel function of degree 3 and cells of size [10, 10] px, best fits the requirements for implementation into the low-cost automated real-time passenger counting system. This is in agreement with an earlier study of passenger recognition without the height differential [19].

The well-established ConvNets such as AlexNet and VGG-16 are expected to be a good basis of object recognition systems. As the validation results (Figure 10) show, they feature good learning ability, resulting in small loss function values. A similar ability can be observed for the AlexNet. From this perspective, the proposed networks Net1 and Net2 seem to be insufficient. However, their classification performance evaluated on the testing set (Table 3) is comparable with AlexNet- and LeNet-5-based systems (there is no clear winner among these four networks). Surprisingly, the VGG-16-based system has the worst performance in the category of the ConvNet-based systems. The most likely explanation of this phenomenon is a relatively high learning capacity of VGG-16 (compared to the other presented architects) that may cause overfitting on the head recognition task. Considering the high time complexity of VGG-16 (Figure 12), we conclude that, despite expectations, VGG-16 is not appropriate for the passenger recognition.

We also investigated possible benefits of the transfer learning by the training of ConvNet-based passenger recognition systems. We observe a lower variability in the cost function values for the networks trained using TL, when compared to the networks trained from scratch (Figure 10). Also, the medians of the cost function values are shifted towards smaller values for TL. We conclude that a model with a low cost function value can be more likely obtained using TL than by its training from scratch.

The size of cells has been reported to be the seminal parameter predetermining the performance of object recognition systems which are based on HOG features [27]. The experimental results presented in this article confirm this finding. An incorrect setting of the cell size results in inferior classification (compare results obtained for cells of size [10, 10] px and [14, 14] px in Figure 11 and Table 3). Also, the time complexity of the HOG-based system strongly depends on the setting of this parameter (compare, e.g., results for cells of size [6, 6] px and [10, 10] px in Figure 12).

5. Conclusions

Presently, deep ConvNets are usually considered as the first choice when developing an image recognition system. We established that image recognition systems with equally good classification performances can be developed using traditional computer vision methods. When appropriately designed and setup, such systems can beat ConvNets-based solutions in terms of time efficiency which is particularly important in real-world applications. This is also the case of the HOG-based passenger recognition system, where the utilization of HOG features in combination with the SVM

classifier can result in time-efficient and accurate passenger recognition. In this context, we showed that passenger heads are sufficient for the precise while fast passenger recognition. We also showed that the HOG-based system is highly flexible, as it can be employed in both low-floor and high-floor means of transport. Its implementation into a passenger monitoring system is being currently developed, allowing us utilization of a basic processing unit. Cost savings on the unit is reflected in the final price of the person flow monitoring system and thus supports its mass use in means of transport all over the world.

Data Availability

Data used to support the findings of the study are available at https://www.researchgate.net/publication/342888989_Dataset_for_head_detector.

Conflicts of Interest

The authors declare that they do not have any commercial or associative interest that represents conflicts of interest in connection with the work submitted.

Acknowledgments

This work was supported from ERDF/ESF “Cooperation in Applied Research between the University of Pardubice and companies, in the Field of Positioning, Detection and Simulation Technology for Transport Systems (PosiTrans)” (no. CZ.02.1.01/0.0/0.0/17_049/0008394).

References

- [1] A. Olivo, G. Maternini, and B. Barabino, “Empirical study on the accuracy and precision of automatic passenger counting in european bus services,” *Open Transportation Journal*, vol. 13, no. 1, pp. 250–260, 2020.
- [2] M. Siebert and D. Ellenberger, “Validation of automatic passenger counting: introducing the T-test-induced equivalence test,” *Transportation*, 2019.
- [3] J. W. Choi, X. Quan, and S. H. Cho, “Bi-directional passing people counting system based on IR-UWB radar sensors,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 512–522, 2018.
- [4] Z. Chen, W. Yuan, M. Yang, C. Wang, and B. Wang, “SVM based people counting method in the corridor scene using a single-layer laser scanner,” in *Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, November 2016.
- [5] S. Akamatsu, N. Shimaji, and T. Tomizawa, “Development of a person counting system using a 3D laser scanner,” in *Proceedings of the 2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*, Bali, Indonesia, December 2014.
- [6] A. Ahmed and N. A. Siddiqui, “Design and implementation of infra-red based computer controlled monitoring system,” in *Proceedings of the 2005 Student Conference on Engineering Sciences and Technology*, Karachi, Pakistan, August 2005.
- [7] T. Teixeira, G. Dublon, and A. Savvides, “A survey of human-sensing: methods for detecting presence, count, location,

- track, and identity,” Technical report, Yale University, New Haven, CT, USA, 2010.
- [8] M. Mohaghegh and Z. Pang, “A four-component people identification and counting system using deep neural network,” in *Proceedings of the 2018 5th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, Nadi, Fiji, December 2018.
 - [9] V. A. Sindagi and V. M. Patel, “A survey of recent advances in CNN-based single image crowd counting and density estimation,” *Pattern Recognition Letters*, vol. 107, 2018.
 - [10] S. Sun, N. Akhtar, H. Song, C. Zhang, J. Li, and A. Mian, “Benchmark data and method for real-time people counting in cluttered scenes using depth sensors,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3599–3612, 2019.
 - [11] L. Del Pizzo, P. Foggia, A. Greco, G. Percannella, and M. Vento, “Counting people by RGB or depth overhead cameras,” *Pattern Recognition Letters*, vol. 81, pp. 41–50, 2016.
 - [12] L. Sun, Y. Liu, S. Chen, B. Luo, Y. Li, and C. Liu, “Pig detection algorithm based on sliding windows and PCA convolution,” *IEEE Access*, vol. 7, pp. 44229–44238, 2019.
 - [13] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, “Selective search for object recognition,” *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2013.
 - [14] W. Liu, D. Anguelov, D. Erhan et al., “SSD: single shot multibox detector,” in *Computer Vision—ECCV 2016*, pp. 21–37, Springer International Publishing, Cham, Switzerland, 2016.
 - [15] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
 - [16] J. Redmon and A. Farhadi, “YOLO9000: better, faster, stronger,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017.
 - [17] X. Wu, “Design of person flow counting and monitoring system based on feature point extraction of optical flow,” in *Proceedings of the 2014 5th International Conference on Intelligent Systems Design and Engineering Applications*, Changsha, China, June 2014.
 - [18] A. Tokta and A. K. Hocaoglu, “A fast people counting method based on optical flow,” in *Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, September 2018.
 - [19] M. Belloc, S. A. Velastin, R. Fernandez, and M. Jara, “Detection of people boarding/alighting a metropolitan train using computer vision,” in *Proceedings of the 9th International Conference on Pattern Recognition Systems (ICPRS 2018)*, Valparaiso, Chile, May 2018.
 - [20] T. Ma, Q. Ji, and N. Li, “Scene invariant crowd counting using multi-scales head detection in video surveillance,” *IET Image Processing*, vol. 12, no. 12, pp. 2258–2263, 2018.
 - [21] M. B. Shami, S. Maqbool, H. Sajid, Y. Ayaz, and S.-C. S. Cheung, “People counting in dense crowd images using sparse head detections,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 9, pp. 2627–2636, 2019.
 - [22] S. I. Cho and S.-J. Kang, “Real-time people counting system for customer movement analysis,” *IEEE Access*, vol. 6, pp. 55264–55272, 2018.
 - [23] J. Lee and B. Al, “Chapter 19—video surveillance,” in *The Essential Guide to Video Processing*, pp. 619–651, Academic Press, Boston, MA, USA, 2009.
 - [24] P. Sharma and A. Singh, “Era of deep neural networks: a review,” in *Proceedings of the 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2017.
 - [25] Y. Xu, X. Zhou, S. Chen, and F. Li, “Deep learning for multiple object tracking: a survey,” *IET Computer Vision*, vol. 13, no. 4, pp. 355–368, 2019.
 - [26] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, Washington, DC, USA, June 2005.
 - [27] P. Škrabánek and F. Majerik, “Detection of grapes in natural environment using HOG features in low resolution images,” *Journal of Physics: Conference Series*, vol. 870, no. 1, Article ID 012004, 2017.
 - [28] D. J. Lope and A. Dolgun, “Measuring the inequality of accessible trams in Melbourne,” *Journal of Transport Geography*, vol. 83, Article ID 102657, 2020.
 - [29] J. Preston, “Big buses in a small country: the prospects for bus services in Wales,” *Research in Transportation Economics*, vol. 59, pp. 379–387, 2016.
 - [30] A. Kathuria, M. Parida, Ch. Ravi Sekhar, and A. Sharma, “A review of bus rapid transit implementation in India,” *Cogent Engineering*, vol. 3, no. 1, 2016.
 - [31] P. Skrabanek and F. Majerik, “Evaluation of performance of grape berry detectors on real-life images,” in *Proceedings of the 22nd International Conference on Soft Computing*, Brno, Czech Republic, 2016.
 - [32] P. Škrabánek and P. Doležel, “Robust grape detector based on SVMs and HOG features,” *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 3478602, 17 pages, 2017.
 - [33] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
 - [34] Basler, “Basler ace,” 2020, <https://www.baslerweb.com/en/products/cameras/area-scan-cameras/ace/aca2500-60uc/>.
 - [35] Computar, “Computar lenses,” 2020, <https://computar.com/product/705/M3514-MP>.
 - [36] NVIDIA, “Nvidia jetson,” 2020, <https://developer.nvidia.com/embedded/jetson-tx2>.
 - [37] N. Aloysius and M. Geetha, “A review on deep convolutional neural networks,” in *Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, April 2017.
 - [38] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
 - [39] L. Bottou, C. Cortes, J. S. Denker et al., “Comparison of classifier methods: a case study in handwritten digit recognition,” in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, Jerusalem, Israel, October 1994.
 - [40] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
 - [41] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
 - [42] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015, <https://arxiv.org/abs/1409.1556>.
 - [43] P. D. Kingma and J. Ba, “Adam: a method for stochastic optimization,” 2014, <https://arxiv.org/abs/1412.6980>.

- [44] ITU-R Recommendation BT 601, *Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide Screen 16:9 Aspect Ratios*, ITU, Geneva, Switzerland, 2011.
- [45] C. H. Lampert, "Kernel methods in computer vision," *Foundations and Trends® in Computer Graphics and Vision*, vol. 4, no. 3, pp. 193–285, 2008.
- [46] Y. Sasaki and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, 2007.
- [47] J.-J. Zhang and P. Zhong, "Learning biased SVM with weighted within-class scatter for imbalanced classification," *Neural Processing Letters*, vol. 51, no. 1, pp. 797–817, 2020.
- [48] P. Dolezel and D. Stursa, "Dataset for head detector," 2020, https://www.researchgate.net/publication/342888989_Dataset_for_head_detector.

Research Article

Heterogenous Trip Distance-Based Route Choice Behavior Analysis Using Real-World Large-Scale Taxi Trajectory Data

Yajuan Deng,¹ Meiye Li,² Qing Tang,³ Renjie He,¹ and Xianbiao Hu ³

¹College of Transportation Engineering, Chang'an University, Xi'an 710064, China

²School of Transportation, Southeast University, Nanjing 211189, China

³Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA

Correspondence should be addressed to Xianbiao Hu; xbhu@mst.edu

Received 4 May 2020; Revised 4 August 2020; Accepted 25 August 2020; Published 9 September 2020

Academic Editor: Kun Xie

Copyright © 2020 Yajuan Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most early research on route choice behavior analysis relied on the data collected from the stated preference survey or through small-scale experiments. This manuscript focused on the understanding of commuters' route choice behavior based on the massive amount of trajectory data collected from occupied taxicabs. The underlying assumption was that travel behavior of occupied taxi drivers can be considered as no different than the well-experienced commuters. To this end, the DBSCAN algorithm and Akaike information criterion (AIC) were first used to classify trips into different categories based on the trip length. Next, a total of 9 explanatory variables were defined to describe the route choice behavior, and the path size (PS) logit model was then built, which avoided the invalid assumption of independence of irrelevant alternatives (IIA) in the commonly seen multinomial logit (MNL) model. The taxi trajectory data from over 11,000 taxicabs in Xi'an, China, with 40 million trajectory records each day were used in the case study. The results confirmed that commuters' route choice behavior are heterogenous for trips with varying distances and that considering such heterogeneity in the modeling process would better explain commuters' route choice behaviors, when compared with the traditional MNL model.

1. Introduction

Analysis of the routing choice behavior provides theoretical support for route guidance and traffic assignment. Most early research studies on route choice behavior were based on the data collected from stated preference (SP) surveys or through small-scale experiments that were usually limited in data size or number of participants. In the modeling process, discrete choice models especially logit models were commonly used. The differences among these models were mainly reflected in the differences of data set, explanatory variables, or the model structure. For example, McFadden and Reid applied logit models to travel behavior analysis [1]. After that, based on the hypothesis that the random term of route utility function follows the Gumbel distribution, Dial constructed a discrete multinomial logit (MNL) model for multimode selection [2, 3]. In order to address the

independence of irrelevant alternatives (IIA) issue of the MNL model, various modified models were proposed, such as the C-logit model and PS-logit model [4, 5], which were built by adding a modification term in the utility function to characterize the interactions among different routes. In addition, according to the generalized extreme value (GEV) theorem proposed by McFadden, some researchers proposed CNL and PCL models [6, 7] to avoid the IIA assumption of the MNL model. In general, these early research studies on route choice behavior lacked real-world data and were restricted by the algorithm complexity, and the numbers of explanatory variables used were usually limited as well.

With the rapid advancement of information and communication technologies (ICT), GPS technology has made significant progress, and the data collected by GPS devices have been widely used in various transportation research,

such as in travel time estimation [8–10], driving risk analysis [11, 12], departure time modeling [13, 14], and many others [15–17]. Such data have also been used to directly support the route choice behavior analysis, and the data-driven route choice models were qualitatively improved in terms of both effectiveness and accuracy. For example, route choice behaviors and network information in Chicago were studied using data collected using portable GPS devices, and path size (PS) logit models for different travel purposes in different time periods were proposed [18]. Based on the same method, Schussler and Axhausen collected travel data in Zurich area and calibrated C-logit model and PS-logit model [19]. Kim Mahmassani proposed a trajectory clustering algorithm to analyze the spatial and temporal travel patterns in a network [20], in which a framework for clustering and classifying vehicle trajectory data was built. Additionally, several medium-sized cities in Netherlands were selected as research objects and an MNL model based on GPS data was proposed to analyze the route choice behavior [21]. Li et al. collected the GPS data of private cars in Toyota City, explored the effect of travelers' heterogeneity on route choice, and concluded that the route choice behavior is affected by travelers' age, gender, vehicle displacement, and O-D's characteristics [22]. However, the analysis focus was on the traveler's heterogeneity, as opposed to the differences on the route characteristics. Bierlaire and Frejinger used the GPS data in Swiss to study the behavior characteristics of long-distance travel route selection and gave the estimation results of the PS-logit model and subnetwork model [23]. Miwa et al. used the taxi travel data of Nagoya City to analyze the characteristics of dynamic route choice behavior, an MNL model was built, and it was concluded that there are differences in the route choice behavior at different O-D distances [24]. Yamamoto et al. used the pedestrian GPS data from Nagoya to build a nested logit model [25], and Hu et al. used GPS data to analyze route choice behavior changes under preplanned road closures [26].

This manuscript focused on the analysis of route choice behavior of general traffic, based on the massive amount of trajectory data collected from the occupied taxicabs. Taxicabs, especially those work with the e-hailing platform such as Uber and Lyft, on the other hand, are mostly installed with the GPS devices for dispatching and safety purposes. However, most existing research studies based on the taxi GPS trajectory data focused on the routing behavior of the vacant taxi drivers, with the objective of minimizing the search time for the next customer [27, 28] or maximizing the profit [29–31], which was significantly different from regular drivers. Our underlying assumption was that when a taxi was occupied by customers, the taxi driver would seek to arrive at the destination in the least amount of time or distance as expected or required by the customer, similar to the objective of a commuter in his/her own car. Additionally, taxi drivers usually had good knowledge on the roadway network and traffic conditions, and thus their travel behavior can be considered as very similar to, and no different than, the well-experienced commuters.

Furthermore, this manuscript tested a hypothesis that trips with different lengths may exhibit different

characteristics in driver's route choice behavior. As opposed to the common practice of developing and calibrating a unified model to describe the route choice behavior of all trips, the Akaike information criterion (AIC) was first used to classify trips into different categories based on the trip length. Next, a total of 9 explanatory variables were defined to describe the route choice behavior, and a PS-logit model was then built, which avoided the invalid assumption of IIA in the commonly seen multinomial logit model [24]. The taxi trajectory data from over 11,000 taxicabs in Xi'an, China, with 40 million trajectory records each day were used in the case study. The results confirmed the hypothesis that commuters' route choice behaviors are heterogeneous for trips with varying distances and that considering such heterogeneity in the modeling process would better explain commuters' route choice behaviors.

The rest of this paper is organized as follows: Section 2 presents the data used in this research, including the GPS trajectory data and the traffic network. Section 3 discusses the analysis methodology in depth, and Section 4 presents the numerical analysis results. Section 5 concludes this research.

2. Data Preparation

2.1. GPS Data Set. The GPS trajectory data used in this research came from the historical database of the taxi dispatch system in Xi'an, China. The recording time was from 0:00 to 24:00, the recording interval was 30 s, and each record contained license plate number, timestamp, longitude, latitude, speed, driving direction, and loading state. The data set included data from over 11,000 taxicabs with 40 million trajectory records each day. Such a huge amount of data can meet the needs of this research. The following data cleaning and preprocessing were performed:

- (1) Removed the flawed data with missing values.
- (2) Only kept the data with loading state being "5 (passenger)."
- (3) Removed the data with driving direction beyond 0° – 360° .
- (4) Removed the data with key attributes being "0 (invalid)."

2.2. Traffic Network. The OpenStreetMap (OSM) network of Xi'an was downloaded and utilized for this research. Post-processing efforts were made, including the removing the duplicate or redundant roads and adding the length of road segment and node information. Additionally, the road segments were classified into seven categories, including expressway, national highway, other highways, urban expressway, main road, secondary road, and neighborhood street. The research region is shown in Figure 1.

2.3. Hotspot OD Trips Extraction. Occupied trips between frequent origin–destination (OD) pairs were extracted from the database as the target data for analysis. We first identified

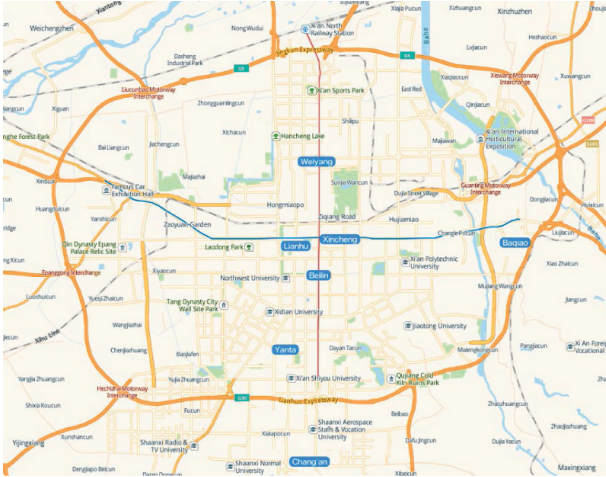


FIGURE 1: Traffic network in Xi'an city.

pick-up and drop-off hotspots and then extracted the frequent OD between these hotspots.

2.3.1. Identification of Drop-Off Hotspots. This step aimed to identify the areas with high density of drop-off events, with the goal of providing the basis for hotspot OD matching and ensuring that there was a sufficient number of passenger-carrying trips between the same OD pair (from pick-up to drop-off).

According to the change of loading state between two adjacent GPS data records, pick-up points and drop-off points can be identified. Taking GPS data of taxi on 19 April 2017 in Xi'an as an example, from 40 million trajectory data generated by 11,281 taxis, nearly 594 thousand drop-off points in the research region were obtained. The DBSCAN spatial clustering algorithm [32] was adopted to identify the drop-off hotspots. The algorithm contained two parameters: cluster neighborhood radius (Eps) and minimum density threshold (MinPts). In this paper, the K -distance method was used to determine the reasonable Eps. The method contained three steps:

Step 1: assuming that the drop-off points data set $D = \{P_i(x, y), i = 1, 2, \dots, n\}$ contained n points, we selected a drop-off point $P_i(x, y)$ and calculated the Euclidean distances between $P_i(x, y)$ and $P_1(x, y), P_2(x, y), \dots, P_{i-1}(x, y), P_{i+1}(x, y), \dots, P_n(x, y)$, respectively. Then, they were sorted by Euclidean distances in ascending order as $d_i^1, d_i^2, \dots, d_i^{k-1}, d_i^{k+1}, \dots, d_i^n$ in which d_i^k indicated the K -distance of drop-off point $P_i(x, y)$.

Step 2: we calculated the K -distance of each drop-off point in the data set based on Step 1.

Step 3: we sorted the K -distances of all drop-off points in ascending order and plotted the K -distance figure. In the figure, the K -distance of the inflection point was defined as Eps of the data set.

Taking the drop-off points data set on Wednesday, April 19, 2017, as an example, we analyzed the data in different

lengths of time. We found that when the length of the time period exceeds 8 minutes, the change of the K -distance figure tends to be stable, and the characteristics of the inflection point are more clear, which was shown in Figure 2. Finally, considering the limitations of computer performance, we took the drop-off points data set (5,000–5,400 points in total) of 10:00–10:10 am on Wednesday, April 19, 2017, as an example, and its K -distance figure is shown in Figure 2(d), which showed the K -distance changed significantly around 0.00211. Therefore, 0.00211 was selected as the Eps. This value will be used in the clustering of one day's drop-off points data set to identify the hotspot ODs.

MinPts indicated the density of drop-off points in each cluster. In this paper, with the given Eps and assuming MinPts, clustering results of drop-off points can be obtained. According to the clustering results under different MinPts, reasonable MinPts can be determined. Under different MinPts, the clustering results of the drop-off points are shown in Table 1.

To obtain as many clusters as possible and to ensure each cluster has a sufficient number of pick-up or drop-off points, the value of MinPts was set to be 800. The 594 thousand drop-off points were clustered into 11 clusters (Table 1). When the value of MinPts was set to be 800, spatial distribution of drop-off clusters and number of trips of each cluster were obtained as shown in Figure 3.

2.3.2. Identification of Hotspot OD. In order to ensure that the trip between the selected OD pairs is of sufficient quantity and effectiveness, a hotspot OD identification method was proposed in this step. It consisted of the following two steps: (1) For each drop-off point in drop-off clusters shown in Figure 3 (14283 points in total), search the corresponding pick-up point and trajectory data in between; (2) Re-cluster the pick-up points. The DBSCAN algorithm was used for the re-clustering of pick-up points. The pick-up points generated by the 11 drop-off clusters, as shown in Figure 3, were re-clustered. Eighteen pairs of hotspot ODs were obtained (Table 2). The results show that using the method above only needs to process one day's data to ensure that the number of passenger-carrying routes between ODs is sufficient.

In Table 2, CCluster means a pick-up hotspot that was re-clustered. "Cluster 1–CCluster:245" means there are 245 single passenger-carrying trips between the pick-up point Cluster1 and the drop-off point CCluster.

3. Analysis Methodology

3.1. Trip Length Classification. To test the hypothesis of heterogeneous route choice behavior for trips with different lengths, the Akaike information criterion (AIC) was first used to classify trips into different categories based on their length.

A few studies on the classification of trips by travel distance can be found in the literature. In the survey of urban residents' travel, the travel distance was subjectively divided into few distance segments, such as 0~3 km, 3~6 km,

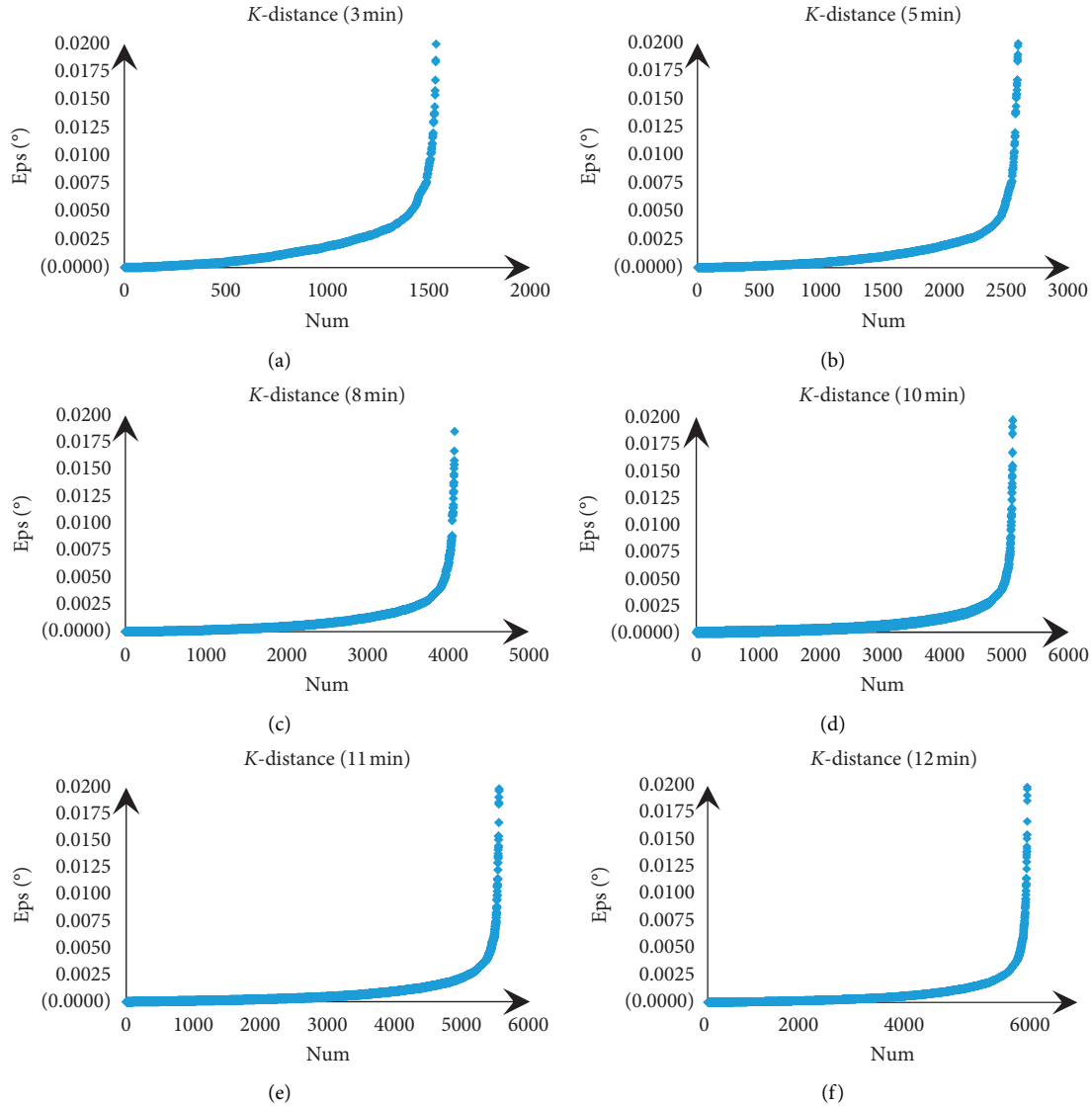


FIGURE 2: K-distance graph of GPS data in different time periods: (a) 3 min, (b) 5 min, (c) 8 min, (d) 10 min, (e) 11 min, and (f) 12 min.

TABLE 1: Clustering results of the pick-up (drop-off) points.

MinPts	Cluster of drop-off points	Noise data	Proportion of noise data	Cluster of pick-up points	Noise data	Proportion of noise data (%)
650	35	135277	68.03	34	135137	68.02
700	24	158980	79.95	26	158839	79.95
750	20	174013	87.51	17	173819	87.49
800	11	183478	92.27	11	183316	92.27
850	6	190179	95.64	6	190031	95.65
900	4	191671	96.39	3	191521	96.40

6~9 km, 9~12 km, and longer than 12 km [33, 34]. For mode split purpose, only qualitative classification of travel distance (short distance and long distance) was performed [35, 36]. In the route choice model, most studies used only one model to describe all the route choice behaviors [8, 21, 37]. For different types of passenger-carrying routes, the behavior of travelers was different. As such, currently a theoretical-sound method for classifying the travel routes is missing.

Based on the OD-Euclidean distance distribution of passenger-carrying routes, we sought for the eigenvalues with the travel volume changes significantly. These eigenvalues were used as the basis for the preliminary classification. The OD-Euclidean distance distribution of the 14,283 trips in 11 drop-off clusters mentioned in Section 2.3 is shown in Figure 4. In this section, we use this part of data for research.

Cluster	Cluster 1	Cluster 2	Cluster 3
Number of trips	914	2950	1004
Cluster	Cluster 4	Cluster 5	Cluster 6
Number of trips	977	1508	1338
Cluster	Cluster 7	Cluster 8	Cluster 9
Number of trips	979	855	1077
Cluster	Cluster 10	Cluster 11	Sum
Number of trips	1880	802	14283

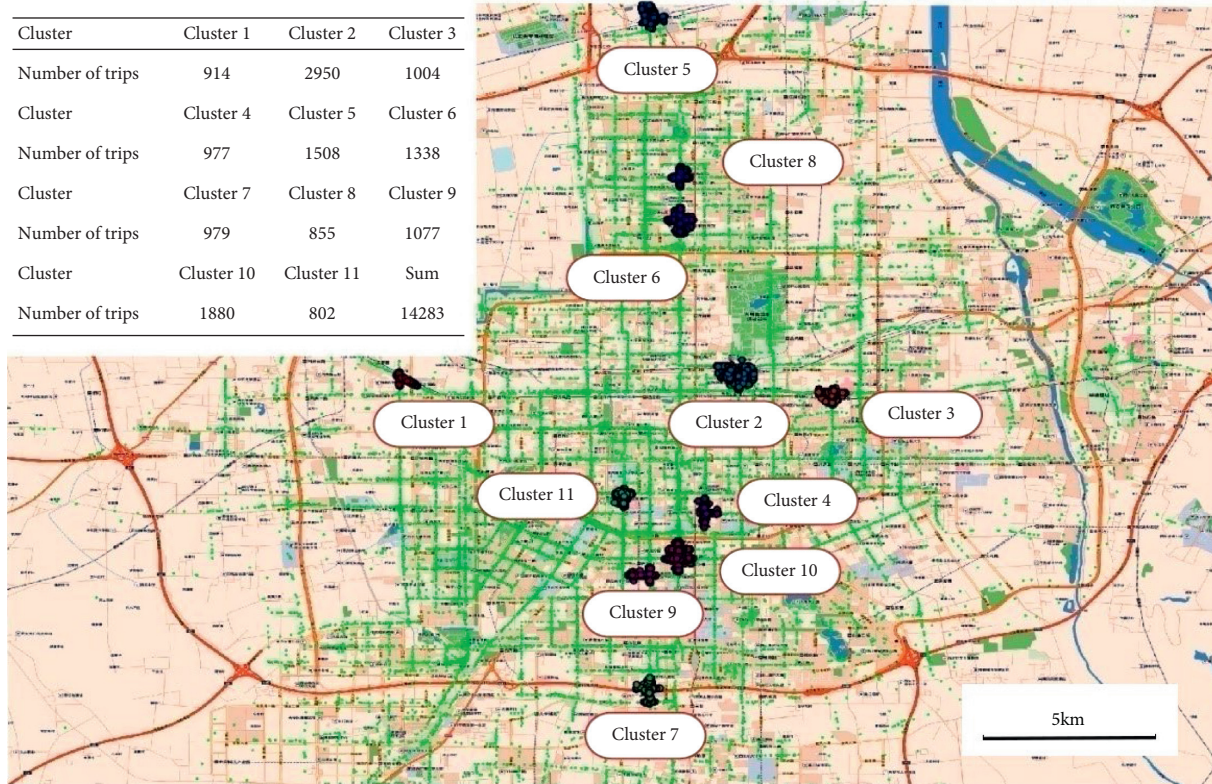
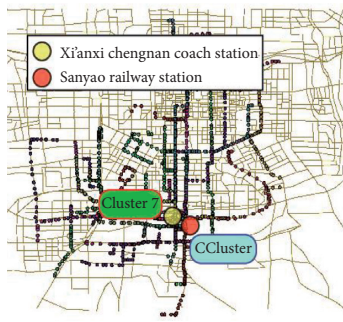


FIGURE 3: Spatial distribution of drop-off clusters.

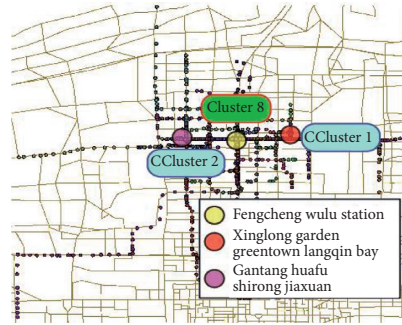
TABLE 2: Results of hotspot OD matching.



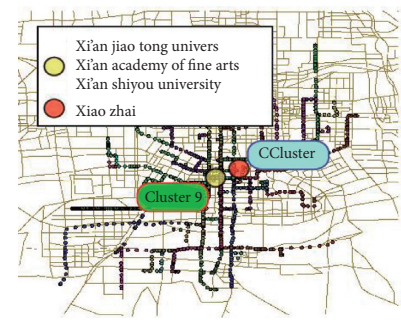
TABLE 2: Continued.



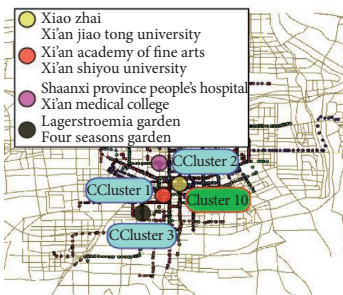
(10) Cluster 7–CCluster : 327



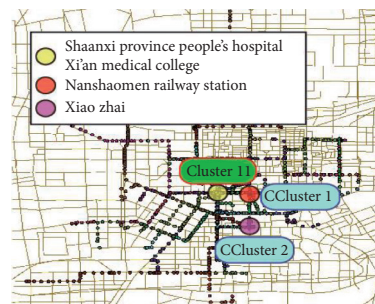
(11) Cluster 8–CCluster 1:137
(12) Cluster 8–CCluster 2:267



(13) Cluster 9–CCluster : 109



(14) Cluster 10–CCluster 1 : 306
(15) Cluster 10–CCluster 2 : 193
(16) Cluster 10–CCluster 3 : 176



(17) Cluster 11–CCluster 1 : 137
(18) Cluster 11–CCluster 2 : 204

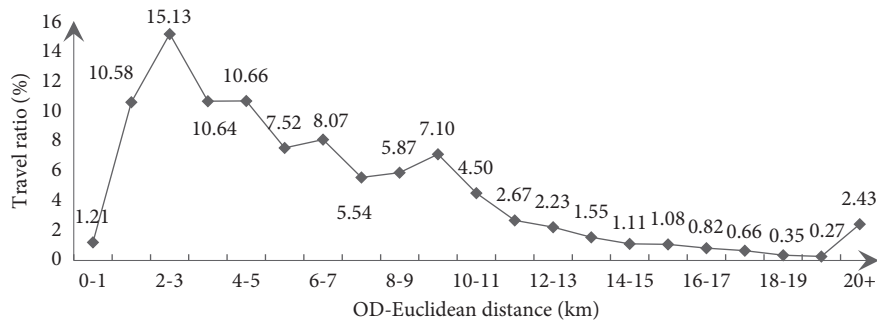


FIGURE 4: OD-Euclidean distance distribution of 11 drop-off clusters.

Figure 4 shows that at 3, 7, and 10 km, three peak values of travel volume can be observed. It is believed that these three peaks were consistent with the urban structure of Xi'an:

- (1) 3 km radius: within 1–3 km of the Central Business District (CBD), there were many service facilities. These facilities can serve residents well and residents can fulfill their daily needs in this region, such as working, schooling, and shopping.
- (2) 7 km radius: as a city with thousands of years of history, the CBD of Xi'an attracted a large number of trips. The CBD of Xi'an is located in the geometric

center of the city, and CBD-centered 6-7 km covered major urban areas.

- (3) 10 km radius: there are many passenger stations, airports, and tourist areas around the city, and these important points of interests also attracted a lot of travel. This phenomenon explains the occurrence of the third peak.

According to the above analysis, single passenger-carrying route of taxi can be divided into four categories: 0–3 km, 3–7, 7–10 km, and longer than 10 km. It should be noted that these were OD-Euclidean distances, which represented the linear distances between pick-up point and

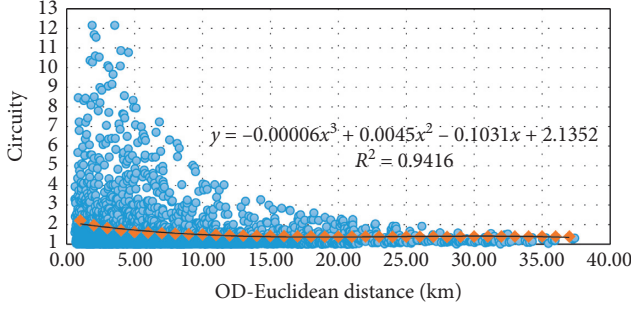


FIGURE 5: OD-Euclidean distances and circuity distribution.

drop-off point. It was difficult to reflect the actual length or travel time of the routes. In order to reflect the actual length of taxi passenger-carrying route, circuity was selected as another route classification index. We screened out the data of 14,283 trips, including the Euclidean distances and circuity of each OD as shown in Figure 5. The relationship between OD-Euclidean distances and average circuity for different types of passenger-carrying routes was fitted as follows. This is a typical regression curve fitting done using Microsoft Excel, and the results show an R square value of 0.9416, which indicates satisfactory results:

$$\text{cir}_{rs} = -0.00006 \cdot \text{eucl}_{rs}^3 + 0.0045 \cdot \text{eucl}_{rs}^2 - 0.1031 \cdot \text{eucl}_{rs} + 2.1352, \quad (1)$$

where cir_{rs} is the circuity of passenger-carrying route from pick-up point r to drop-off point s and was calculated by the ratio of OD-Euclidean distance to the actual travel distance. eucl_{rs} is the OD-Euclidean distance of passenger-carrying route from pick-up point r to drop-off point s (unit: kilometer).

The mean values of 0–3 km, 3–7 km, and 7–10 were 1.5 km, 5 km, and 8.5 km, respectively. Considering that only 13.17% of OD-Euclidean distances were over 10 km, and 80% of them were distributed in 10–15 km, and 12.5 km was selected as the representative value. By introducing 1.5 km, 5 km, 8.5 km, and 12.5 km into equation (1), the initial clustering centers of five schemes can be calculated (1.9905, 1.7247, 1.5471, and 1.4324). In addition, there have been studies that divide the travel distance of travelers into 3 categories and above [35, 36]. Therefore, we decided to set the number of clusters to 3 or 4. If the number of clusters is 3, depending on the cluster center, there are 4 optional clustering schemes; if the number of clusters is 4, there is 1 optional clustering scheme. Five clustering schemes are shown in Table 3.

In order to compare the effect of the five clustering schemes, the AIC criterion, proposed by H. Akaike in information theory, was introduced to identify the best scheme.

$$\text{AIC} = -2 \cdot \ln L(\hat{\theta}) + 2 \cdot k, \quad (2)$$

where $L(\hat{\theta})$ is the maximum likelihood estimation of the model, and with the increase in the difference among clusters, the value becomes larger. k is the number of parameters in the model, and the more classifications the

TABLE 3: Initial clustering centers of five clustering schemes.

Scheme	Number of clusters	Initial clustering centers			
1		1.9905	1.7247	1.5471	
2	3	1.9905	1.7247	1.4324	
3		1.9905	1.5471	1.4324	—
4		1.7247	1.5471	1.4324	
5	4	1.9905	1.7247	1.5471	1.4324

model consists, the greater the value will be. The value of AIC depends on $L(\hat{\theta})$ and k . The smaller the k is, the more concise the model becomes, and the larger the $L(\hat{\theta})$ is, the more accurate the model will be. The AIC therefore considered both complexity and precision in identifying the best scheme.

For circuity data sets $\{X_i | i = 1, 2, \dots, K\}$, which contained K circutries of passenger-carrying routes. The number of clusters was N , the final cluster center of each cluster was $\{C_m | m = 1, 2, \dots, N\}$, sample size of each cluster was $\{Q_m | m = 1, 2, \dots, N\}$, and internal deviation of each cluster was $\{D_m | m = 1, 2, \dots, N\}$.

$$D_m = \frac{1}{Q_m} \cdot \sum_{j=1}^{Q_m} \text{dist}(X_j, C_m), \quad (3)$$

where $\text{dist}(X_j, C_m)$ is the Euclidean distance between X_j and C_m , X_j is the circuity of a passenger-carrying route in cluster m , and C_m is the center of cluster m .

The density distribution of deviations in each cluster is shown in equation (4).

$$f(D_m) = \frac{Q_m/K}{(d_{\max} - d_{\min})/N} = \frac{N}{K} \cdot \frac{Q_m}{d_{\max} - d_{\min}}, \quad (4)$$

where $d_{\max} = \text{Max}(D_m)$ and $d_{\min} = \text{Min}(D_m)$.

According to the principle of logarithmic maximum likelihood estimation, the logarithmic maximum likelihood estimation functions of the internal deviations of each cluster ($\{D_m | m = 1, 2, \dots, N\}$) can be obtained as follows:

$$\ln L(D_m | m = 1, 2, \dots, N) = -N \cdot \ln \frac{K}{N} - \sum_{m=1}^N \ln \frac{d_{\max} - d_{\min}}{Q_m}. \quad (5)$$

Plug equation (5) into equation (2), the AIC, which was the basis of passenger-carrying route classification, can be calculated as follows:

$$\text{AIC}(N) = 2 \sum_{m=1}^N \ln \frac{d_{\max} - d_{\min}}{Q_m} + 2N \left(1 + \ln \frac{K}{N} \right). \quad (6)$$

The clustering scheme with minimum AIC were selected as the optimal scheme. Five K -means clustering schemes were implemented by SPSS, which is a statistical analysis software package developed by IBM, and the AIC values of the five schemes, which as shown in Table 3, were 2.885, 2.6137, 2.8041, 3.5233, and 3.0231, respectively. The AIC value of scheme 2 was the smallest, which means that this scheme had the best balance in complexity and precision.

Accordingly, scheme 2 was considered as the optimal scheme.

In clustering scheme 2, the boundaries of cluster 1 were 1 and 1.489, which corresponded to the passenger-carrying routes with OD-Euclidean distance longer than 10 km. The boundaries of cluster 2 were 1.489 and 1.826, which corresponded to the passenger-carrying routes with OD-Euclidean distance between 3 km and 10 km. The boundaries of cluster 3 were 1.826 and 2.544, which corresponded to the passenger-carrying routes with OD-Euclidean distance between 0 km and 3 km. Accordingly, the classification results of taxi passenger-carrying routes were $0 \text{ km} \leq D \leq 3 \text{ km}$ (short distance), $3 \text{ km} \leq D \leq 10 \text{ km}$ (medium distance), and $10 \text{ km} < D$ (long distance), where D indicated the OD-Euclidean distance.

With such thresholds for trip lengths clarification, the Euclidean distance distribution of 18 pairs of hotspot ODs is shown in Figure 6.

The hotspot OD from Xiaozhai (Cluster18, pick-up cluster) to Shaanxi Province People's Hospital and Xi'an Medical College (Cluster11, drop-off cluster) was selected as the research object of short-distance taxi passenger-carrying route. The hotspot OD from Lagerstroemia Garden and Four Seasons Garden (Cluster16, pick-up cluster) to Xiaozhai (Cluster10, drop-off cluster) was selected as the research object of medium-distance taxi passenger-carrying route. The hotspot OD from Xi'an Bei Railway Station (Cluster2, pick-up cluster) to Xi'an Railway Station (Cluster2, drop-off cluster) was selected as the research object of long-distance taxi passenger-carrying route. These three OD pairs are illustrated in Figure 7.

3.2. Route Choice Probability Distribution Analysis. Figure 8 illustrates the actual probability distribution of route choice for different passenger-carrying route categories shown in Figure 7. The formula for calculating the fluctuation value of the path choice probability is as follow:

$$\Delta P = \max_k P_{ik}^{rs} - \min_k P_{ik}^{rs}, \quad (7)$$

where P_{ik}^{rs} stands for the probability of driver i choosing route k taxi from r to s .

It can be observed that the fluctuation of route choice probability can be summarized as follow: 0.2010 (short distance) < 0.239 (long distance) < 0.305 (medium distance). The following can be found:

- (1) *Short-distance passenger-carrying routes* had the smallest fluctuation. A most likely explanation was that due to the limited scale of the network between short-distance hotspot OD pair, drivers did not have enough options to make a detour and utility values of difference routes were similar.
- (2) *Medium-distance passenger-carrying routes* had the highest fluctuation. The scale of network between medium-distance hotspot OD pair was moderate, as drivers had more options to make a detour in acceptable travel time.

- (3) The fluctuation of *long-distance passenger-carrying routes* was higher than short-distance routes but lower than medium-distance routes. It was probably because that the scale of network between long-distance hotspot OD pair was large and drivers had enough options to make a detour. However, the drivers' acceptable circuitry or delays were small for long-distance passenger-carrying routes.

3.3. Explanatory Variables. In this study, route choice behavior modeling explanatory variables were selected from three aspects: path factor, road factor, and PS correction term. We defined the coefficients corresponding to the explanatory variables in the model as shown in Table 4 below.

In Table 4, the travel time (TT) equals to the difference between the origin and destination GPS timestamps of a single passenger-carrying trip, K represents the length of path, D represents the OD-Euclidean distance, N_p is the number of intersections, K_m stands for the length of main road, K_s represents the length of secondary, K_b represents the length of branch road, and K_{co} is the length of congested road, which is judged by the average travel speed of the road section from GPS data.

3.4. Path Size Logit Model. The traditional multinomial logit model was a discrete choice model based on the theory of random utility, which can be used to describe the individual's choice behavior. The model was simple and easy to understand. However, the IID assumption of utility random item led to the result that there were IIA characteristics in the model. The probability that two routes were selected was only related to the utility of them and not to other routes. However, according to Figure 6, we knew that there were many common roadway segments among different taxi passenger-carrying routes.

The path size logit model reflected this issue by introducing a correction term into the utility function. Therefore, the PS-logit model was adopted to analyze the taxi passenger-carrying route choice behavior in this paper. The utility function of PS-logit is shown in equation (8).

$$U_{ik}^{rs} = V_{ik}^{rs} + \beta_{PS}^{rs} \ln(\text{PS}_k^{rs}) + \varepsilon_k^{rs}. \quad (8)$$

U_{ik}^{rs} : utility of traveler i choosing route k , from pick-up point r to drop-off point s .

V_{ik}^{rs} : fixed utility of traveler i choosing route k , from pick-up point r to drop-off point s .

β_{PS}^{rs} : parameters to be calibrated.

PS_k^{rs} : path-size value of route k , from pick-up point r to drop-off point s .

$$\text{PS}_k^{rs} = \sum_{a \in \Gamma_k} \frac{l_a}{L_k^{rs}} \cdot \frac{1}{\sum_{j \in K_{rs}} \mu_{aj}^{rs}}, \quad \forall k \in K_{rs}, rs \in \text{RS}. \quad (9)$$

Γ_k : roads set in route k .

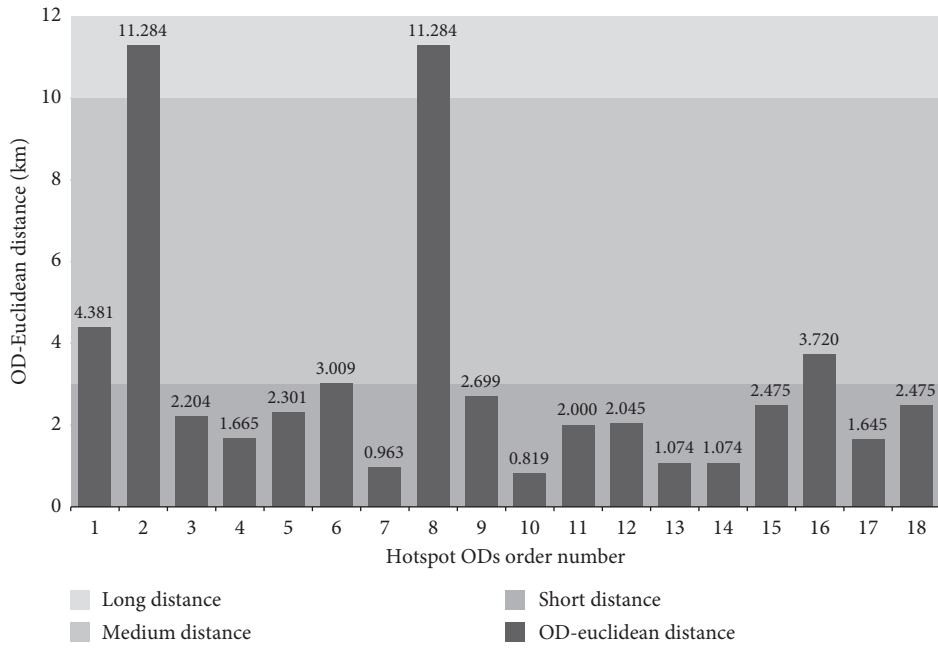


FIGURE 6: Euclidean distance distribution of hotspot ODs.

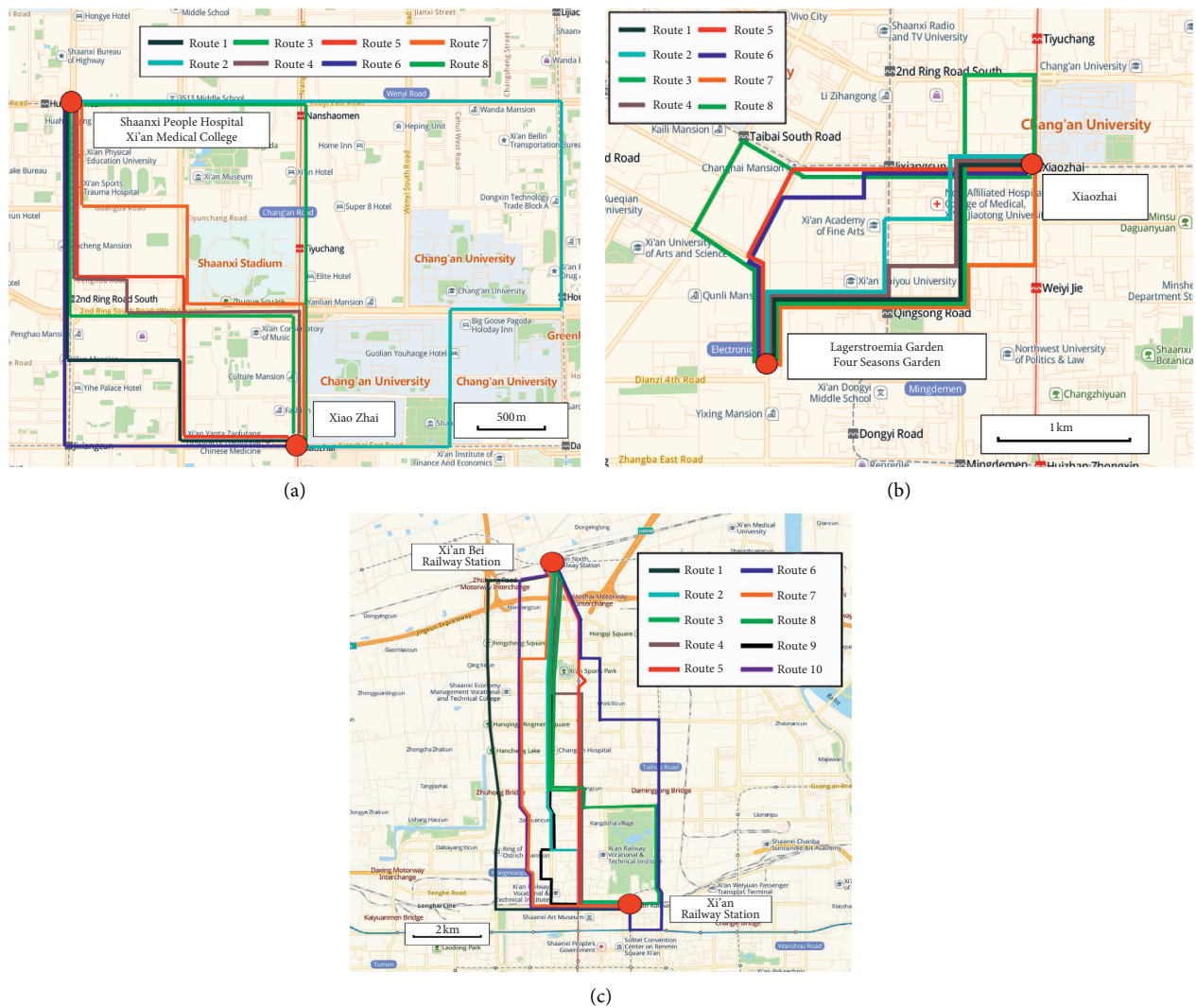


FIGURE 7: Routes of different passenger-carrying route categories: (a) short-distance passenger-carrying routes, (b) medium-distance passenger-carrying routes, and (c) long-distance passenger-carrying routes.

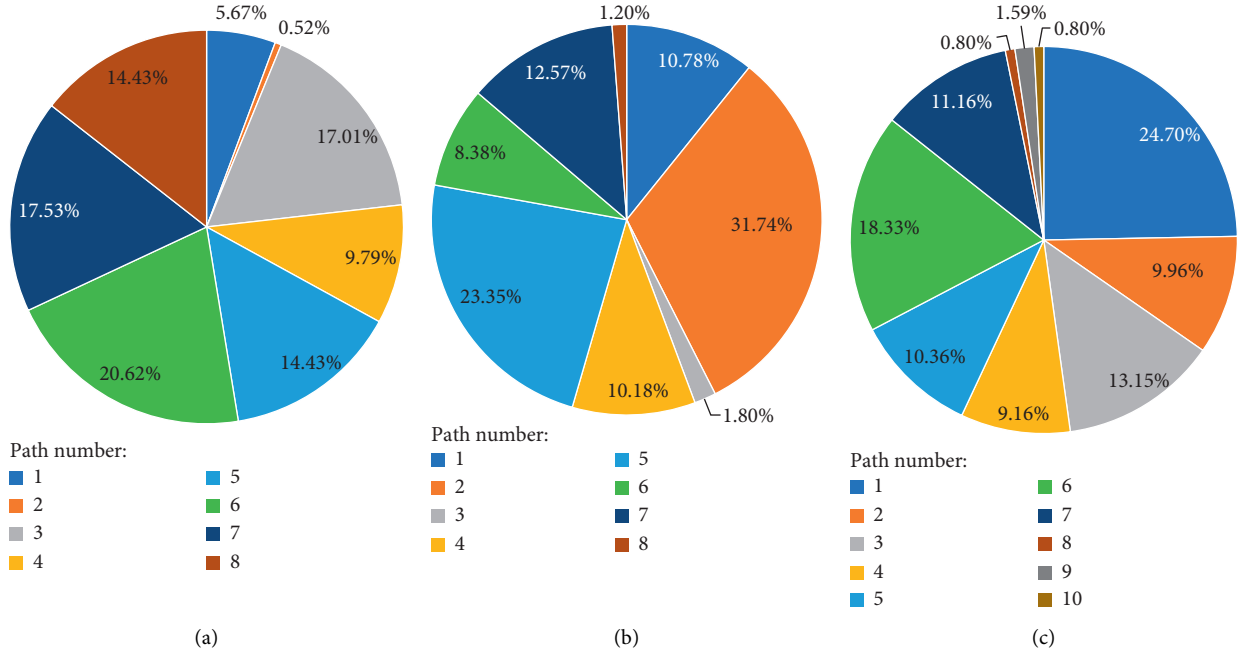


FIGURE 8: Probability distribution of route choice for different route categories: (a) short distance, (b) medium distance, and (c) long distance.

TABLE 4: Explanatory variables of taxi passenger-carrying route choice behavior.

Variables	Abbreviation	Coefficient	Unit	Equation
Route factors	Travel time	TT	Minute	—
	Circuitry	CI	—	K/D
	Frequency of intersections	PT	—/minute	N_p/TT
	Proportion of main road	TR	%	K_m/K
Road factors	Proportion of secondary road	SR	%	K_s/K
	Proportion of branch road	BR	%	K_b/K
	Left turns	L	—	—
	Right turns	R	—	—
	Proportion of congestion	CO	—	K_{co}/K
Path size	Ln(PS)	Ln(PS)	—	—

K_{rs} : routes set between r (pick-up point) and s (drop-off point).

RS: OD set.

l_a : length of road a .

L_k^{rs} : length of route k .

μ_{aj}^{rs} : if road a belongs to route j , μ_{aj}^{rs} equals to 1, otherwise μ_{aj}^{rs} equals to 0.

The PSL model for this study is constructed as follows:

$$P_{ik}^{rs} = \frac{\exp\left[\sum_{m=1}^M (\beta_m \cdot c_{m,ik}^{rs}) + \beta_{PS}^{rs} \cdot \ln(\text{PS}_k^{rs})\right]}{\sum_{l \in K_{rs}} \exp\left[\sum_{m=1}^M (\beta_m \cdot c_{m,il}^{rs}) + \beta_{PS}^{rs} \cdot \ln(\text{PS}_l^{rs})\right]} \quad (10)$$

P_{ik}^{rs} : for the taxi passenger-carrying route from r to s , the probability that driver i chooses route k .

β_m : coefficient of explanatory variable m .

$c_{m,ik}^{rs}$: for the taxi passenger-carrying route from r to s , when the driver chooses route k , the value of explanatory variable m .

M : the number of explanatory variables.

4. Results and Discussion

4.1. Model Calibration Results. With the help of Biogeme software package, the parameters of MNL model and PS-Logit model with different types of passenger-carrying routes were calibrated, respectively. In addition, we aggregate all routes together as a control group. The results are shown in Table 5.

According to Table 5, for different route types, the t-statistics of explanatory parameters of the two models were statistically valid. The coefficient of PS correction term was positive, which was consistent with the basic principle of the

TABLE 5: Calibration results of taxi passenger-carrying route choice model.

Variable	Short distance				Medium distance				Long distance				All trips combined			
	PSL		MNL		PSL		MNL		PSL		MNL		PSL		MNL	
	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat	Coef	<i>t</i> -stat
β_{TT}	-0.41	-4.21	-0.61	-5.84	-0.46	-3.34	-0.62	-3.89	-0.55	-2.81	-0.83	-3.28	-0.74	-2.19	-0.86	-3.91
β_{CI}	-2.35	-2.62	-3.3	-3.42	-2.45	-2.88	-3.7	-3.92	-3.45	-2.2	-4.76	-3.77	-4.72	-2.71	-4.32	-2.66
β_{PT}	0.02	2.48	0.5	3.41	0.02	2.05	0.55	2.86	0.04	1.97	0.79	2.71	0.13	2.34	-0.12	3.26
β_{TR}	0.02	1.78	0.53	1.91	0.01	1.76	0.52	1.84	0.03	1.96	0.7	2.03	0.03	1.80	0.67	2.04
β_{SR}	-0.01	-1.69	-0.02	-1.83	-0.02	-1.88	-0.02	-1.9	-0.02	-1.78	-0.02	-1.7	-0.02	-1.72	-0.08	-2.49
β_{BR}	-3.87	-2.01	-5.43	-2.78	-8.29	-1.87	-7.83	-2.91	-22.35	-2.31	-24.37	-3.21	-10.10	-2.77	-7.54	-2.26
β_L	-0.42	-3.36	-0.53	-4.54	-0.45	-3.7	-0.54	-5.41	-0.58	-5.66	-0.82	-7.16	-0.47	-3.35	-0.19	-1.77
β_R	-0.17	-2.54	-0.43	-2.69	-0.17	-2.56	-0.45	-2.73	-0.21	-1.76	-0.48	-2.52	-0.38	-1.73	-0.43	-3.21
β_{CO}	-1.59	-3.07	-1.61	-3.56	-1.85	-2.29	-1.88	-2.58	-2.38	-3.56	-2.41	-3.17	-2.68	-2.06	-3.22	-2.01
Ln (PS)	3.47	3.2	—	—	4.02	3.69	—	—	4.71	5.11	—	—	3.37	2.12	—	—
Num of observations	204				176				262				642			
Likelihood ratio ρ^2	0.272		0.225		0.271		0.223		0.316		0.197		0.219		0.182	

TABLE 6: MRS of explanatory variables.

Factors	Symbol	MRS		
		Short distance	Medium distance	Long distance
Circuitry	$MRS(\beta_{CI}, \beta_{TT})$	5.73170	5.32609	6.27272
Frequency of intersections	$MRS(\beta_{PT}, \beta_{TT})$	-0.04878	-0.04348	-0.07273
Proportion of main road	$MRS(\beta_{TR}, \beta_{TT})$	-0.04878	-0.02174	-0.05455
Proportion of secondary road	$MRS(\beta_{SR}, \beta_{TT})$	0.02439	0.04348	0.03636
Proportion of branch road	$MRS(\beta_{BR}, \beta_{TT})$	9.43902	18.02174	40.63636
Left turns	$MRS(\beta_L, \beta_{TT})$	1.02439	0.97826	1.05454
Right turns	$MRS(\beta_R, \beta_{TT})$	0.41463	0.36957	0.38182
Proportion of congestion	$MRS(\beta_{CO}, \beta_{TT})$	3.87805	4.02174	4.32727
Path size	$MRS(\text{Ln}(\text{PS}), \beta_{TT})$	-8.46341	-8.73913	-8.56364

PS-logit model. In addition, adjusted likelihood ratio of PS-Logit model was better than that of the MNL model, which meant that PS-logit model described drivers' passenger-carrying route choice behavior more accurately than the traditional MNL model. Finally, the adjusted likelihood of the control group was significantly lower than the other three groups, which showed that dividing the passenger-carrying route by distance can optimize the model. According to Table 5, the following conclusions can be drawn:

- (1) The coefficients with positive values included β_{PT} , β_{TR} , and Ln (PS). The coefficients with negative values included β_{TT} , β_{CI} , β_{SR} , β_{BR} , β_L , β_R , and β_{CO} . This showed that when drivers chose routes, they tended to choose roads with high proportion of main roads, lower circuitry, shorter travel time, and less congestion, regardless of the length of travel distance.
- (2) With the increase of travel distance, the absolute value of β_{CI} , β_{PT} , β_{TR} , β_{SR} , β_{BR} , and β_{CO} increased obviously. This indicated that as travel distance increases, the impacts of circuitry, path structure, and the congestion proportion of the choice of the driver will also increase.

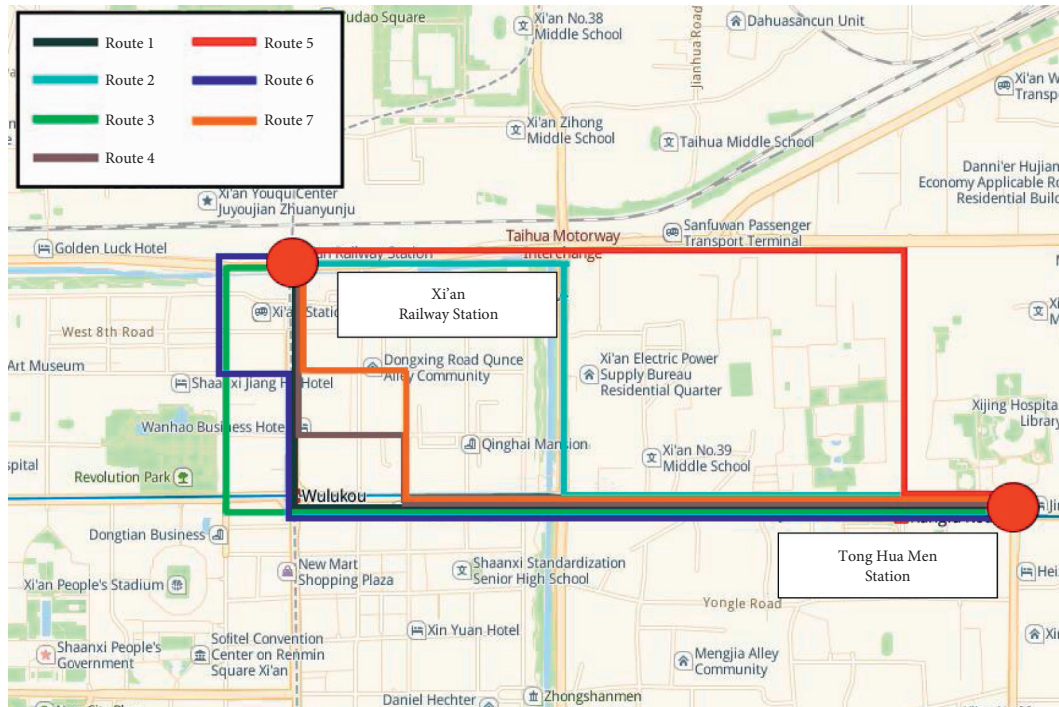
4.2. *Route Choice Preference Analysis.* With the level of consumer satisfaction unchanged, the marginal rate of substitution (MRS) referred to the scenario that when consumers increased one unit of a product and needed to abandon certain number of another product. Many existing research studies use MRS in the analysis of the calibration results of the choice model [38, 39]. In this paper, with the utility of passenger-carrying route kept unchanged, MRS was defined as the change of basic variable when the other explanatory variables increased by one unit. It can be calculated as follows:

$$MRS(\alpha_i, \alpha_j) = \frac{\partial U / \partial \alpha_i}{\partial U / \partial \alpha_j} \quad (11)$$

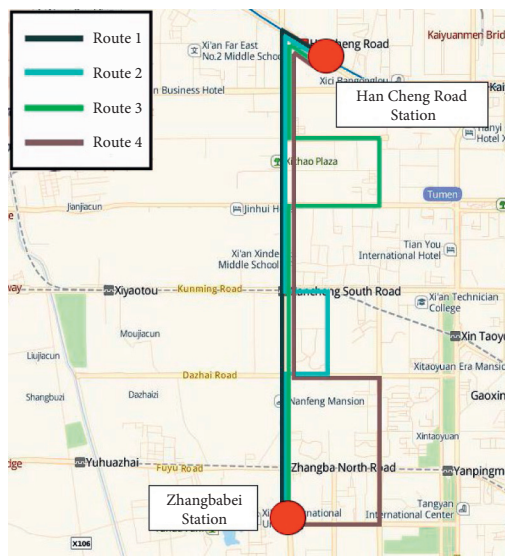
In this study, the PS-Logit model with a better adjusted likelihood ratio was selected as the analysis object. Travel time was selected as the basic variable, the MRS between travel time and other explanatory variables are shown in Table 6.

According to Table 6, the following conclusions can be drawn:

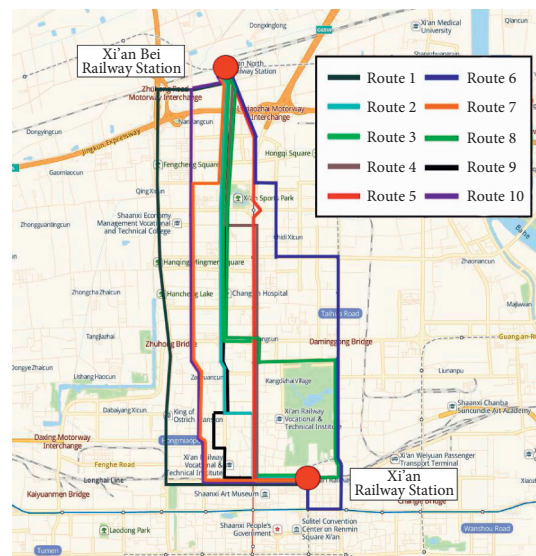
- (1) The relationship among the MRS of explanatory variables was found to be $MRS(\beta_{BR}, \beta_{TT}) > MRS(\text{Ln}(\text{PS}), \beta_{TT}) > MRS(\beta_{CI}, \beta_{TT}) > MRS(\beta_{CO}, \beta_{TT}) > MRS(\beta_L, \beta_{TT}) > MRS(\beta_R, \beta_{TT}) > MRS(\beta_{PT}, \beta_{TT}) > MRS$



(a)



(b)



(c)

FIGURE 9: Routes of different passenger-carrying route categories for verification: (a) short-distance passenger-carrying routes, (b) medium-distance passenger-carrying routes, and (c) long-distance passenger-carrying routes.

TABLE 7: Short-distance passenger-carrying route choice model verification results.

Route number	Actual choice times	True choice ratio	Calculated choice times	Calculated choice ratio	Number of errors
Route 1	233	0.524	213	0.479	20
Route 2	79	0.178	90	0.203	11
Route 3	66	0.148	47	0.105	19
Route 4	23	0.052	41	0.092	18
Route 5	16	0.036	23	0.052	7
Route 6	11	0.025	8	0.018	3
Route 7	17	0.038	23	0.052	6
Total	445	1.000	445	1.000	84
Hit ratio			0.81421		

TABLE 8: Medium-distance passenger-carrying route choice model verification results.

Route number	Actual choice times	True choice ratio	Calculated choice times	Calculated choice ratio	Number of errors
Route 1	154	0.815	136	0.720	18
Route 2	12	0.063	19	0.101	7
Route 3	9	0.048	5	0.026	4
Route 4	14	0.074	29	0.153	15
Total	189	1.000	189	1.000	44
Hit ratio			0.76720		

TABLE 9: Long-distance passenger-carrying route choice model verification results.

Route number	Actual choice times	True choice ratio	Calculated choice times	Calculated choice ratio	Number of errors
Route 1	77	0.266	71	0.247	6
Route 2	32	0.111	29	0.100	3
Route 3	35	0.121	38	0.131	3
Route 4	22	0.076	26	0.092	4
Route 5	25	0.087	31	0.104	5
Route 6	57	0.197	53	0.183	4
Route 7	27	0.093	32	0.112	5
Route 8	3	0.010	2	0.008	1
Route 9	7	0.024	5	0.016	2
Route 10	4	0.014	2	0.008	2
Total	289	1.000	289	1.000	35
Hit ratio			0.87889		

$(\beta_{TR}, \beta_{TT}) > MRS(\beta_{SR}, \beta_{TT})$. If the goal was to reduce travel time, the first and foremost factors to be considered should be proportion of branch road, path-size value, circuitry, and proportion of congestion. The minor factors to be considered should be the number of left turn, right turn, number of nodes per minute, and the proportion of main road and secondary road.

- (2) As the distance of passenger-carrying route increased, the MRS of circuitry and proportion of branch road and congestion also increased. On the contrary, the MRS of frequency of intersections decreased. When the distance of passenger-carrying route was long, drivers usually avoided routes with high circuitry and proportion of congestion and preferred to choose the routes with high proportion of freeway or highway segments.
- (3) To maintain the utility of passenger-carrying route unchanged, if the number of left turn increased by one, for short-distance, medium-distance, and long-distance passenger-carrying routes, the travel time needed to be reduced by 1.02, 0.98, and 1.05 min, respectively. If the number of right turns increased by one, for short-distance, medium-distance, and long-distance passenger-carrying routes, the travel time needed to be reduced by 0.41, 0.37, and 0.38 min, respectively. Time cost of left turn was about 2.6 times as high as that of right turn.

4.3. Verification of Route Choice Model. The verification of the path selection model is mainly achieved by comparing the trial calculation results of the path selection model with the actual path selection results, and finally the model's hit

ratio to evaluate the effectiveness of the model is obtained. The calculation steps of hit ratio are as follows:

Step 1: assuming that the total number of samples is N , the total number of alternatives is M , there are K parameters in the final calibration result of the model, and the parameter calibration value β_k and the corresponding parameter value C_k are brought into the calibration model to obtain the selection probability \hat{P}_m of the corresponding program.

Step 2: assuming that traveler n has the greatest probability of choosing the route m , then $\hat{\delta}_{nm} = 1$, otherwise $\hat{\delta}_{nm} = 0$.

Step 3: when the actual selection result δ_{nm} of the traveler is consistent with the predicted result of the calibration model, set $S_m^n = 1$, otherwise $S_m^n = 0$. Then, the hit rate can be calculated as follows:

$$HR = \frac{1}{N} \sum_{n=1}^N S_m^n. \quad (12)$$

In this paper, three different types of OD, as shown in Table 2, were selected to verify the model: The hotspot OD from Tong Hua Men Station (CCluster3, pick-up cluster) to Xi'an Railway Station (Cluster2, drop-off cluster) was selected as the verification object of short-distance taxi passenger-carrying routes; the hotspot OD from Han Cheng Road Station (CCluster, pick-up cluster) to Zhangbabei Station (Cluster1, drop-off cluster) was selected as the verification object of medium-distance taxi passenger-carrying routes; and the hotspot OD from Xi'an Railway Station (Cluster2, pick-up cluster) to Xi'an Bei

Railway Station (CCluster2, drop-off cluster) was selected as the verification object of long-distance taxi passenger-carrying routes. After removing abnormal data, these three ODs have 445, 189, and 289 valid trips and 7, 4, and 10 valid routes, respectively. The routes between these three ODs are shown in Figure 9.

According to the route choice model constructed in Section 4.1, the route choice results of each hotspot OD are calculated and compared with the actual choice situation. The results are shown in Tables 7–9.

The Tables 7–9 show that the hit ratios of the short-distance, medium-distance, and long-distance passenger-carrying route choice models are 0.81421, 0.76720, and 0.87889, respectively, indicating that the three types of route choice models constructed are effective and can be explained reasonably the behavior of passenger-carrying route choice. The analysis of extra OD pairs requires the substantial amount of manual work.

5. Conclusion and Future Work

This manuscript, for the first time, focused on the analysis of route choice behavior based on the massive amount of real-world GPS trajectory data collected from the occupied taxi cabs. Our analysis based on the trajectory data from Xi'an, China, found that for trips with different lengths, the characteristics of route choice behavior could be very different. As such, according to the distribution of Euclidean distance and volume, five route classification schemes for taxi passenger-carrying routes were proposed based on the circuitry *K*-means clustering method. The Akaike information criterion (AIC) principle was adopted to identify the best route classification scheme. After that, taxi passenger-carrying routes were divided into three categories: short distance, medium distance, and long distance. Based on the MNL model, three PS-Logit models were proposed to analyze the route choice behaviors. The numerical analysis validated our hypothesis and revealed heterogenous activity patterns and influencing factors for trips with different lengths.

According to the study, the following conclusions can be drawn: (1) taxi passenger-carrying routes can be classified based on the distribution of Euclidean distance and *K*-means clustering of circuitries; (2) for different taxi passenger-carrying routes, the fluctuation of route choice probability can be summarized as follows: short distance < long distance < medium distance; (3) for different taxi passenger-carrying routes, the first and foremost factors to be considered were proportion of branch road, path-size value, circuitry, and proportion of congestion. The minor factors to be considered were the number of left turns, right turns, the number of nodes per minute, and the proportion of main road and secondary road; (4) with the increase of travel distance, drivers usually avoided routes with high circuitry and intersection density but preferred to choose the routes with high proportion of freeway or highway; and (5) the effects of circuitry, frequency of intersections, path structure, and congestion degree on utility function were significantly

different among different taxi passenger-carrying route categories.

Finally, we have selected another OD pair for each category for validation purpose, and the analysis shows consistent conclusions. Future research could be focused on using the data set from other cities to validate the model. The works to be improved are as follows: On the one hand, the variables considered in the model in this paper were easy to be defined, while some other factors that were difficult to be defined or computed were not taken into account such as trip purpose, preference, network familiarity, and influence of weather and environment. On the other hand, in this manuscript, only Euclidean distance, travel volume, and circuitry were considered in the taxi passenger-carrying route classification. If more data types become available, more factors could be considered such as the network structure among the hotspot OD. How to identify and select sufficient factors to improve the route classification results may need further discussion.

Data Availability

The GPS trajectory data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The research is supported by the National Key Research and Development Program of China (grant no. 2018YFB1600900), the Shaanxi Provincial Science and Technological Project (grant nos. 2020JM-244), and the Science and Technological Project of Shaanxi Provincial Transport Department (grant no. 19-24X).

References

- [1] D. Mcfadden and F. Reid, "Aggregate travel demand forecasting from disaggregate behavioral models," *Transportation Research Record*, vol. 534, pp. 24–37, 1975.
- [2] R. B. Dail, "A probabilistic multipath traffic assignment model, which obviates path enumeration," *Transportation Research*, vol. 5, no. 2, pp. 83–111, 1971.
- [3] M. Ben-Akiva and S. R. Lerman, "Disaggregate travel and mobility choice models and measures of accessibility," in *Proceedings of the 3rd International Conference on Behavioral Travel Modelling*, Tanunda, Australia, 1977.
- [4] E. Cascetta, A. Nuzzolo, F. Russo, and F. Vitetta, "A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks," in *Proceedings of the Thirteenth International Symposium on Transportation and Traffic Theory*, pp. 697–711, Lyon, France, 1996.
- [5] M. S. Ramming, *Network Knowledge and Route Choice*, Massachusetts Institute of Technology, Cambridge, MA, USA, 2002.

- [6] C.-H. Wen and F. S. Koppelman, "The generalized nested logit model," *Transportation Research Part B: Methodological*, vol. 35, no. 7, pp. 627–641, 2001.
- [7] F. S. Koppelman and C.-H. Wen, "The paired combinatorial logit model: properties, estimation and application," *Transportation Research Part B: Methodological*, vol. 34, no. 2, pp. 75–89, 2000.
- [8] Q. Tang and X. Hu, "Modeling individual travel time with back propagation neural network approach for advanced traveler information systems," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 6, Article ID 04020039, 2020.
- [9] Z. Li, R. Kluger, X. Hu, Y.-J. Wu, and X. Zhu, "Reconstructing vehicle trajectories to support travel time estimation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 42, pp. 148–158, 2018.
- [10] H. Noh, A. Sun, X. Hu, and A. Rehan, "Development of a regional network performance measurement model for planning application based on high-frequency GPS data," in *Proceedings of the Transportation Research Board 96th Annual Meeting*, Washington, DC, USA, January 2017.
- [11] X. Zhu, Y. Yuan, X. Hu, Y.-C. Chiu, and Y.-L. Ma, "A Bayesian network model for contextual versus non-contextual driving behavior assessment," *Transportation Research Part C: Emerging Technologies*, vol. 81, pp. 172–187, 2017.
- [12] X. Hu, Y.-C. Chiu, Y.-L. Ma, and L. Zhu, "Studying driving risk factors using multi-source mobile computing data," *International Journal of Transportation Science and Technology*, vol. 4, no. 3, pp. 295–312, 2015.
- [13] X. Hu, X. Zhu, Y.-C. Chiu, and Q. Tang, "Will information and incentive affect traveler's day-to-day departure time decisions?—An empirical study of decision making evolution process," *International Journal of Sustainable Transportation*, vol. 14, no. 6, pp. 403–412, 2020.
- [14] H. Xianbiao, C. Yi-Chang, and Z. Lei, "Behavior insights for an incentive-based active demand management platform," *International Journal of Transportation Science and Technology*, vol. 4, no. 2, pp. 119–134, 2015.
- [15] H. Qi and X. Hu, "Real-time headway state identification and saturation flow rate estimation: a hidden Markov chain model," *Transportmetrica A: Transport Science*, vol. 16, no. 3, pp. 840–864, 2020.
- [16] Y. Deng, X. Luo, X. Hu, Y. Ma, and R. Ma, "Modeling and prediction of bus operation states for bunching analysis," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 9, Article ID 04020106, 2020.
- [17] Y.-J. Deng, X.-H. Liu, X. Hu, and M. Zhang, "Reduce bus bunching with a real-time speed control algorithm considering heterogeneous roadway conditions and intersection delays," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 7, Article ID 04020048, 2020.
- [18] N. S. Dhakar and S. Srinivasan, "Route choice modeling using GPS-based travel surveys," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2413, no. 1, pp. 65–73, 2014.
- [19] N. Schussler and K. W. Axhausen, "Accounting for route overlap in urban and sub-urban route choice decisions derived from GPS observations," in *Proceedings of the 12th International Conference on Travel Behavior Research*, Jaipur, India, 2009.
- [20] J. Kim and H. S. Mahmassani, "Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories," *Transportation Research Part C: Emerging Technologies*, vol. 59, pp. 375–390, 2015.
- [21] T. Thomas and B. Tutert, "Route choice behavior in a radial structured urban network: do people choose the orbital or the route through the city center?" *Journal of Transport Geography*, vol. 48, pp. 85–95, 2015.
- [22] D. Li, T. Miwa, T. Morikawa, and P. Liu, "Incorporating observed and unobserved heterogeneity in route choice analysis with sampled choice sets," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 31–46, 2016.
- [23] M. Bierlaire and E. Frejinger, "Route choice modeling with network-free data," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 2, pp. 187–198, 2008.
- [24] T. Miwa, T. Morikawa, and S. Kurauchi, "Preliminary analysis on dynamic route choice behavior using probe-vehicle data," *Infrastructure Planning Review*, vol. 22, pp. 467–476, 2005.
- [25] T. Yamamoto, S. Takamura, and T. Morikawa, "Structured random walk parameter for heterogeneity in trip distance on modeling pedestrian route choice behavior at downtown area," *Travel Behaviour and Society*, vol. 11, pp. 93–100, 2018.
- [26] X. Hu, Y. Yuan, X. Zhu, H. Yang, and K. Xie, "Behavioral responses to pre-planned road capacity reduction based on smartphone GPS trajectory data: a functional data analysis approach," *Journal of Intelligent Transportation Systems*, vol. 23, no. 2, pp. 133–143, 2019.
- [27] X. Hu, S. Gao, Y. Chiu, and D. Lin, "Modeling routing behavior for vacant taxicabs in urban traffic networks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2284, no. 1, pp. 81–88, 2012.
- [28] Q. Tang, X. Hu, and H. Qi, "Modeling routing behavior learning process for vacant taxis in a congested urban traffic network," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 6, Article ID 04020043, 2020.
- [29] X. Yu, S. Gao, X. Hu, and H. Park, "A Markov decision process approach to vacant taxi routing with e-hailing," *Transportation Research Part B: Methodological*, vol. 121, pp. 114–134, 2019.
- [30] X. Yu, S. Gao, X. Hu, and H. Park, "Multi-cycle optimal taxi routing with e-hailing," in *Proceedings of the Transportation Research Board 97th Annual Meeting*, Washington, DC, USA, 2018.
- [31] X. Yu, S. Gao, and X. Hu, "Optimizing vacant taxis' routing decisions: model-based and model-free approaches," in *Proceedings of the Transportation Research Board Annual Conference*, Washington, DC, USA, 2019.
- [32] T. N. Tran, K. Drab, and M. Daszykowski, "Revised DBSCAN algorithm to cluster data with dense adjacent clusters," *Chemometrics and Intelligent Laboratory Systems*, vol. 120, pp. 92–96, 2013.
- [33] G. Wallner, S. Kriglstein, E. Chung, and S. A. Kashfi, "Visualisation of trip chaining behaviour and mode choice using household travel survey data," *Public Transport*, vol. 10, no. 3, pp. 427–453, 2018.
- [34] Y. Sun, J. Shi, and P. M. Schonfeld, "Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro," *Public Transport*, vol. 8, no. 3, pp. 341–363, 2016.
- [35] W. Zhou, W. Wang, and Z. Guo, "Travel distance of different traffic modes based on different grades of roads," *Journal of Wuhan University of Technology (Transportation Science & Engineering)*, vol. 33, no. 5, pp. 976–979, 2009.
- [36] M. Li, G. Song, and C. Ying, "Research on excessive short distance car trips in urban area," *Journal of Beijing Jiaotong University*, vol. 38, no. 3, pp. 15–21, 2014.
- [37] Y. Yang, E. Yao, and L. Pan, "Taxi route choice behavior modeling based on GPS data," *Journal of Transportation*

- Systems Engineering and Information Technology*, vol. 15, no. 1, pp. 81–86, 2015.
- [38] R. Sebastián, J. C. MunOz, and L. D. Grange, “A topological route choice model for metro,” *Transportation Research Part A: Policy and Practice*, vol. 45, no. 2, pp. 138–147, 2011.
- [39] H. Jeffrey, S. Elizabeth, and C. Billy, “A GPS-based bicycle route choice model for San Francisco, California,” *Transportation Letters*, vol. 3, no. 1, pp. 63–75, 2011.

Research Article

The Importance of Cognitive and Mental Factors on Prediction of Job Performance in Chinese High-Speed Railway Dispatchers

Zizheng Guo,^{1,2,3} Jiaming Zou,^{1,2} Chuanning He,⁴ Xi Tan,^{1,2} Chongshuang Chen,⁵ and Guo Feng⁶ 

¹School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China

²National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 610031, China

³National Engineering Laboratory for Comprehensive Transportation Big Data Application Technology, National Development and Reform Commission, Beijing 100000, China

⁴Dispatch Office of China Railway Chengdu Group Co. Ltd., Chengdu 610082, China

⁵School of Mathematics, Southwest Jiaotong University, Chengdu 611756, China

⁶Psychological Research and Counseling Center, Southwest Jiaotong University, Chengdu 611756, China

Correspondence should be addressed to Guo Feng; fengguo@swjtu.edu.cn

Received 8 January 2020; Revised 20 May 2020; Accepted 29 July 2020; Published 14 August 2020

Academic Editor: Kun Xie

Copyright © 2020 Zizheng Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High-speed railway (HSR) transportation poses a serious challenge to dispatchers, whose job performance plays a critical role for the safety and efficiency of the transportation system. This study examined the relationship between cognitive and mental factors and job performance among Chinese high-speed railway dispatchers and established a predictive model for the job performance of dispatchers. Cognitive abilities involved in train dispatching including working memory and multiobject tracking and potential related mental factors including depression, anxiety, perceived stress, and social support were examined. Job performances were measured by both subjective and objective indicators, i.e., the overall evaluation by supervisors and the delay time by dispatching simulator. Stepwise regression results showed that both cognitive abilities and 2 mental factors (depression and perceived stress) have strong relations with job performance, and tremendous distinction between groups of good and poor performance of HSR dispatchers is revealed. The predictive model accounted for 91% of the overall variance in objective performance indicator and has 96% distinguished accuracy of good and poor groups. These findings imply that cognitive and mental factors should be of great concern to the current practice of Chinese HSR dispatcher selection and management.

1. Introduction

With the share of China accounting for about 2/3 of the world high-speed railway (HSR) mileage [1], the daily transportation of China's high-speed trains is about 1.33 million, which occupies 25.7 percent of the overall passenger traffic [2]. The HSR dispatchers take full efficiency and safety responsibility in conducting the high-speed trains' running in the network. Railway Accident (2011) reports that in the Yong-wen railway, a severe accident of two trains' collision resulted in 40 deaths and 172 injuries partly due to the HSR dispatcher's overlook of the malfunctioned signalling

equipment and his failure to deliver the correct movement authority to the trains [3]. Although aided by many automated equipment, the safe and efficient dispatching tasks still heavily rely on the HSR dispatchers' multiple capabilities and mental states in emergency situations [4, 5].

Previous research studies on job performance of other occupations have listed possible impact factors: (1) at the level of organization, including the work environment, assessment system, incentive mechanism, and corresponding training [6, 7]; (2) at the individual level, including but not limited to job-related knowledge, professional skills, capabilities, stress, and emotional state [8, 9]. These impact

factors have different predictive effects according to the variety of occupations. As an emerging occupation, HSR dispatching has its own particularity. Hence, we discussed the predictive factors of HSR dispatchers' job performance from the perspective of work analysis.

As the central nervous system of train operation, HSR dispatcher is not only responsible for daily adjustment of train operation, dispatching order issuing, tracking, and controlling the train route but also needs to dispose all kinds of emergencies quickly at any time. Highly complex and important work content, which reduces the prediction of job performance of ordinary tests concentrated on job-related knowledge, professional skills, and capabilities [10, 11], makes the establishment of strict selection and training system as a must [12–14]. The predictive validity of the cognitive ability depends on the complexity of the job with the strongest validity coefficients observed for highly complex jobs [15, 16]. Literature [17, 18] and task analysis of Chinese HSR dispatchers both reveal that train dispatching is a highly cognitively demanding task [19]. We first summarized HSR dispatchers' job responsibilities and the related cognitive abilities in Table 1.

Among cognitive abilities, working memory and multiobject tracking caught our considerable attention given their highly frequent requirements in HSR dispatcher's job. Working memory can be defined as a temporary retention of representations that were just recently experienced (but no longer exist in the environment) or recalled from long-term memory [20]. Working memory ability is viewed as one of the most important psychological constructs extensively investigated in the cognitive psychology [21–23] and is more highly correlated with the job performance than others constructs like the intelligence [24]. Working memory has been proved to be a better predictor than others, especially under the similar job cognitive requirements situation of air traffic control system [25]. N-back task is often employed to measure the working memory ability [26, 27].

The multiobject tracking ability is also repeatedly reported to be important in facilitating job efficiency by improving the adequate situational awareness on performance from monitoring tasks and multitasking activities [28, 29]. The multiobject tracking ability is viewed as an important factor for air traffic controllers' selection and training [30]. The multiobject tracking (MOT) task is used generally to examine the multiobject tracking ability of various professional capabilities [31, 32]. As seen from the job responsibilities of the HSR dispatchers, the multiobject tracking abilities are also a suitable factor to predict the job performance. Because many tasks such as routing, monitoring, and managing emergency require searching and tracking multisource information on the control panel, visual tracking of dynamic objects is an important component of the HSR dispatching. In conclusion, we assume that the multiobject tracking ability should be included into the cognitive abilities set.

In addition to cognitive abilities, a series of studies showed that mental health also impacts job performance. In the field of traffic safety, drivers with mental problems have much worse driving performance in both simulator

and road tests, for depression and anxiety could slow down reaction speed and increase accident rates [33–35]. Similar results have also been found in nurses and operating workers [36, 37]. Though it has been widely confirmed [33, 36, 38, 39], little is known among HSR dispatchers. The particularities of HSR dispatching make mental factors to be the validity predictors of job performance. For example, HSR dispatchers had to work 12 hours in each shift and their attention and audio-visual perception must to be always highly alert while working. Moreover, the working environment is monotonous and the workplace is completely enclosed. These above unique work characters could easily lead to emotional problems such as depression and anxiety. Depression is the most prevalent emotional problem, and there is no debate about its impact on job performance [40]. Similar to depression, anxiety also plays its key role. While the absence of positive affect is specific to depression, the physiological hyperarousal is specific to anxiety [41]. Anxiety could impair employees' memory ability, cause insomnia, and make them more aggressive when communicating with others [36]. Research investigated that metro dispatchers had worse mental health than other occupation groups in terms of depression and anxiety [42].

Besides, highly cognitive demanding and safety responsibility work also lead to a state of stress. Survey report from Federal Railroad Administration (FRA) has suggested that most dispatchers perceived stress in their daily work [43]. Perceived stress, as an important composition of mental health, is always associated with negative cognitions and negative mood on the one hand [44, 45]. On the other hand, moderate stress is necessary for HSR dispatchers to keep physiological arousal that helps constant vigilance. Hence, perceived stress need to be test as a predictor of HSR dispatchers' performance. Besides, studies have shown that social support can regulate negative emotions [46, 47] and this makes the ability to perceive social support also be a mental health factor. Therefore, factors such as depression, anxiety, perceived stress, and perceived social support were all under consideration as different dimensions of mental health in this study to support the predictive model of HSR dispatcher's job performance.

Meanwhile, cognitive factors may interact with mental factors. Previous studies have shown that mental health affects cognitive ability [48, 49]. Typically, anxiety has a negative relationship with working memory performance [50]; depression could alter reaction time of multiple cognitive abilities including working memory and attention [51]. Vice versa, cognitive ability could also affect mental factors [52, 53]. Related research found that high cognitive performance has been associated with better mental health, e.g., lower risk of psychiatric disorders [54–56]; low working memory capacity leads to an increased susceptibility to anxiety [57].

Therefore, cognitive factors including working memory and multiobject tracking and mental factors such as depression, anxiety, perceived stress, and perceived social support should be taken into account in the job performance model that needs holistic monitoring and comprehensive

TABLE 1: The HSR dispatcher's job responsibilities and the necessary cognitive abilities.

Job responsibilities of HSR dispatchers	Job contents/professional skills	Job-related cognitive abilities
Schedule preparation	Scheduling the routing and movement of trains with the consideration of construction and maintenance plans	Reasoning, planning, and working memory
Operation monitoring	Monitoring train operation, line conditions, and weather conditions	Working memory, multiobject tracking, and attention
Information collection	Communicating with other railway personnel regarding train and track information	Working memory, multiobject tracking, sustained attention, and communication skills
Emergency disposal	Managing emergent events to protect the safety of the public, railway employees, and railway property	Working memory, multiobject tracking, and situation awareness
Data management	Maintaining records of train movements, track assignments, maintenance activities, and other events	Working memory and multiobject tracking

disposition, but to what extent these factors can explain the impact on performance remains unclear. To the authors' best knowledge, this research is one of the first studies to explore the quantifiable relationships between the cognitive and mental factors and the job performance in the context of HSR dispatching, especially on the prediction of the job-related specific cognitive abilities. It would lead to one of the theoretical and practical bases for the HSR dispatchers' candidate selection and personnel managing.

The aim of this article is to get insights into the role of the cognitive abilities and related mental factors in the HSR dispatching environment and to offer a theoretical prediction model on the job performance of the HSR dispatchers. The remaining part of this article is structured as follows. The methodology section explains the sampling techniques, the methods for measurements, and research procedures. The subsequent section is dedicated to the presentation and analysis of the results. The final section draws conclusions along with a discussion.

2. Materials and Methods

The HSR dispatchers were recruited from a branch of China Railway. Cognitive abilities including the multiobject tracking (MOT) and the working memory (WM) of these HSR dispatchers were measured. And mental factors including depression, anxiety, perceived stress, and perceived social support were measured through questionnaire. The conduction of the tests was in a random order across the subjects to control the carry-out effects. Their job performances were measured by both subjective indicator and objective indicator, i.e., the overall evaluation and the delay time.

2.1. Cognitive Ability Tests. Two E-prime 2.0 programmed tests were selected to measure working memory and visual multiobject tracking abilities.

2.1.1. Working Memory: 2-Back Task. One of the most frequently used methods to measure working memory ability is the n-back task [26, 27]. The 2-back task was conducted in our study. Specifically, a series of white letters (font at 48) were presented on the black background; each of

them was shown for 0.5 s and participants were asked to determine whether it is the same as the letter that was presented two items back in the series. If the answer is same, participants were required to press "Enter" key as soon as possible; otherwise, no button should be pressed (Figure 1). The reaction time (RT, ms) and the correct accuracy (ACC) were recorded as the performance of working memory. Each participant did 6 practice trials and completed 176 trials in the formal test.

2.1.2. Multiobject Tracking (MOT) Task. Multiobject tracking (MOT) paradigms [58] are widely used to study the capability of tracking multiple moving objects simultaneously [59]. In each trial (see Figure 2), 16 squares ($1 * 1$ cm) including targets (colored red) and nontargets (colored green) appear at random locations of a $15 * 15$ cm box. The number of targets was randomly determined ranging from one to seven. In the 5 s phase of attentive tracking, all targets turned green just identical to the nontargets first, and all objects started moving. Then, all objects stopped moving, and participants were asked to judge whether one randomly selected object (marked in white color) was a target or not by pressing buttons. Similar to the above test, performance was measured by the reaction time (RT) and the correct accuracy (ACC). Each participant completed 49 formal trials after they have really understood the whole task through practice trials (4–8 trials).

2.2. Psychological Questionnaire

2.2.1. The Zung Self-Rating Depression/Anxiety Scale (SDS and SAS). The SDS and SAS [60] are two widely used scales to quantify depression and anxiety symptoms. Each of them included 20 worded items that rated from 1–4 to assess the level of depression or anxiety, and higher scores correspond to more frequent symptoms. Both SDS and SAS have good reliability and validity [61, 62], and Cronbach's alpha in the current study is 0.757 and 0.82, respectively.

2.2.2. The Perceived Stress Scale (PSS). The PSS [63] is one of the most commonly used instruments to measure stress perception. 14 items in PSS were designed to measure the extent to which one's life is perceived as "unpredictable,

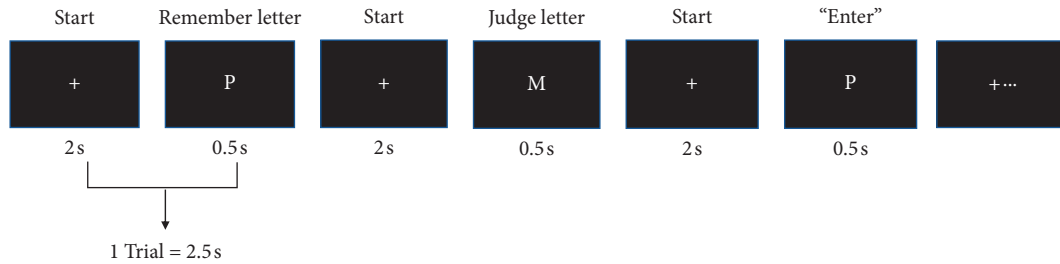


FIGURE 1: Illustration of the 2-back task.

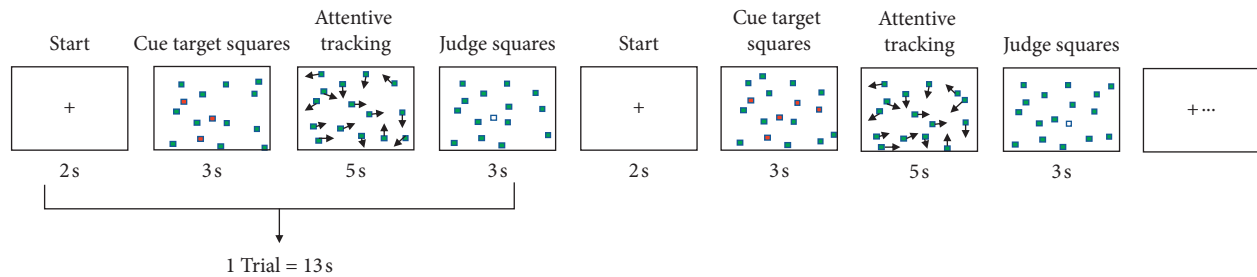


FIGURE 2: Illustration of multiobject tracking (MOT) task.

uncontrollable, and overloading.” Each item was rated on a 5-point Likert scale. Cronbach’s alpha is 0.84 in this study.

2.2.3. The Perceived Social Support Scale (PSSS). The PSSS [64] is the Chinese version of Zimet’s Multidimensional Scale of Perceived Social Support (MSPSS) [65] to measure actually perceived social support. It has 12 items rated on a 7-point Likert scale and is divided into three dimensions: family support, friend support, and other support. Cronbach’s alpha is 0.88 in the current study.

2.3. Evaluation of Job Performance. The HSR dispatcher’s job performance was measured by both subjective indicator (overall evaluation) and objective indicator (delay time). The overall evaluation was made by their supervisors and the delay times were calculated by the HSR dispatching simulator.

2.3.1. Subjective Indicator: Overall Evaluation. The participants’ overall evaluation was given by five supervisors, including two dispatching directors (from two different shifts), a technical training director, a safety director, and an administration director from the HSR dispatching station. The overall evaluation was rated on a 9-point Likert scale (1, completely poor; 2, very poor; 3, poor; 4, rather poor; 5, neutral; 6, rather good; 7, good; 8, very good; 9, completely good) through the question “Please give an overall evaluation of the dispatcher’s job performance.” Before the question, there was a guide paragraph that specified the scoring criteria which should be rated from HSR dispatchers’ five job responsibilities including schedule preparation, operation monitoring, information collection, emergency disposal, and data management. Similar one-item measure

of overall job performance has been used in previous studies [66, 67].

To avoid the inaccurate evaluation from the supervisors who were not familiar with the dispatchers, a 9-point Likert scale question “How well do you know the dispatcher’s job performance of all five job responsibilities?” was also asked. Data from the supervisors with familiarity scores lower than 3 were excluded, which accounted for 3% of all the supervisor-dispatcher pairs. Afterwards, overall evaluation ratings from the remaining supervisors were averaged for each dispatcher.

The Pearson correlation analysis was used to examine the relations between the familiarity and the job performance ($r = 0, p = 1$), which showed that the supervisors’ ratings on the job performance were not likely to be affected by their familiarity with the dispatchers. Also, the r_{WG} values of the job performance ranged from 0.93 to 1 (mean = 0.98; median = 0.98), reaching the cutoff value of 0.7 generally recommended for good inter-rater agreement [68].

2.3.2. Objective Indicator: Delay Time. Delay time is an objective indicator to reflect HSR dispatchers’ job performance. In most circumstances, trains could run normal according to the operation diagram. But they would be stopped or be slowed down when encountering emergency scenarios such as equipment failure, natural conditions, or human causes, which would cause delay for minutes or even hours. Dispatchers need to ensure trains are running on time to the greatest extent when they enter the jurisdiction. Ensuring the running time of trains is an important part of dispatcher’s daily shift and job performance.

To test this objective indicator while controlling all uncertain variables, we use a fully simulated dispatching simulator for data collection. The dispatching simulator was developed by the authors’ team (see Figure 3), which was

entirely designed according to real CTC dispatching system. It was installed with eight Dell computers (Precision 490, Dual Core Intel Xeon Processor 5130 2 GHz) with a 256 MB PCIe \times 16 Nvidia graphic card, Sound Blaster® X-Fi™ system, and Dell A225 Stereo System. A 27-inch LCD screen with 1920 \times 1200 pixel resolution was adopted to display experimental scenarios. The dispatching simulator also has a Huawei tablet that can communicate with other related HSR positions such as drivers. The tablet could respond automatically when the participant spoke. Through this simulator, the dispatcher can conduct emergency dispatching just the same as their real work. Meanwhile, the simulator will record the total delay time of trains in the dispatcher's jurisdiction caused by the dispatcher's operation response and emergency scenario.

2.4. Participants and Procedure. 70 random selected HSR dispatchers from a branch of China Railway participated in this study. All of them were males, which is consistent with the fact that almost all HSR dispatchers in China are males. They were required to avoid alcohol and caffeine on the day of experiment as well as to refrain from eating or exercising two hours before the study. Firstly, participants gave informed consent that participation was voluntary, and it would neither affect their performance evaluation nor have any other consequence in the company. Next, a questionnaire collecting demographic information was conducted, including variables that could be correlated to dispatching skills or cognitive abilities, such as age (mean = 34.76 and SD = 3.25 years), HSR dispatching experience ranging from 1 to 9 years (mean = 4.51 and SD = 2.22 years), and education (82% had a bachelor's or higher degree), and variables that could be possibility correlated to mental factors, such as marriage (89% were married).

Then, all the participants completed 4 mental health tests (integrated into a long questionnaire test) and 2 cognitive tests. The order in which participants took the mental health questionnaire and cognitive tests and the order of WM task and MOT task were all random and counterbalanced. A short break of about three minutes was provided between any two tests (see Figure 4).

Next, the simulator experiment was conducted. In this part, the participants firstly spent several minutes to be familiar with the simulator until they had mastered the basic operation. Secondly, participants were informed of the emergency scenario that needs disposal. The emergency scenario designed by highly experienced HSR dispatcher was the indicator of one of the rail switches turning into a wrong status. According the operation and safety regulations, eight process were included in this situation: (1) an alarm window was popped up on the control panel and the alarm sound was broadcasted; (2) the dispatcher identified the failure and commended the HSR driver to stop the train; (3) the dispatcher confirmed the failure and started the emergency procedure on the control panel; (4) the dispatcher informed the concerned personnel to make preparation; (5) the dispatcher adjusted the train operation plan for safety protection; (6) the station attendant sent a request for track-

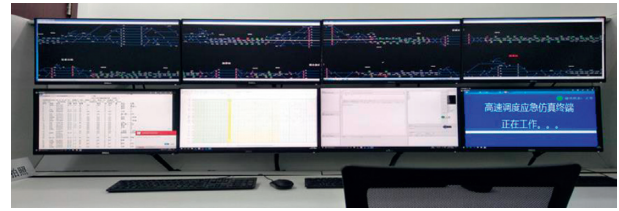


FIGURE 3: The dispatching simulator.

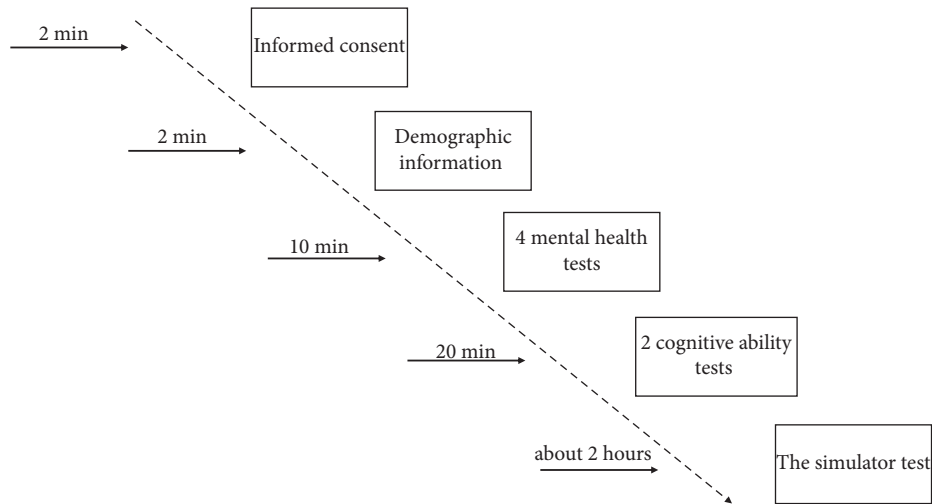
checking to the dispatcher who would issue the related dispatching order; (7) the dispatcher asked the on-site personnel about the situation; (8) the train operation was recovered after the failure was solved. Participants were asked to operate the simulator in actual situation to deal with the failure situation. They did the trial twice including a practice and a formal test. The delay time of all the trains in jurisdiction was obtained from the formal test. Specific operation steps are illustrated in Figure 5.

Participants operated the simulator about 2 hours that depended on their dispatching ability. Each dispatcher participant was given a gift for their participation. At last, 70 questionnaires and simulator data were validated, giving a validity rate of 100%.

3. Results

On basis of the qualitative analysis, this section would explore the analytical relationship between the job performance and the possible variables from a quantitative perspective. As shown in Figure 6, firstly, we divided the dispatchers into the good performance group (Good) and the poor performance group (Poor) and established the threshold for them. Then, in order to establish a predictive model of the dispatcher's job performance, the stepwise regression method was adopted into model of the relation between the HSR dispatchers' job performance and the possible variables. Finally, according to the threshold and the predicted delay time, the dispatchers were identified as Good and Poor. The validity and accuracy of the model would be tested by the comparison between the predicted group and the original group. Specific data analysis procedures are shown in Figure 6.

3.1. Descriptive Statistics of All Variables. In the previous section, the job performance of HSR dispatchers was measured by the subjective and the objective indicators, i.e., the overall evaluation and the delay time. Four demographic variables (age, dispatching experience, education, and marriage), four cognitive variables (ACC of MOT, RT of MOT, ACC of WM, and RT of WM), and four mental variables (anxiety, depression, perceived stress, and social support) were the possible variables to predict the job performance. The mean and standard deviations of these variables are shown in Table 2. Among these demographic variables, "education" and "marriage" were not included in the following statistics for three reasons: (1) these two are dichotomous variables; (2) most of the answers are in one category (82% had a bachelor's or higher degree and 89%



Annotation:
 1. The order of which tests were done first was random and counterbalanced
 2. Half did WM task first and the other half did MOT task first

FIGURE 4: The experimental procedures.

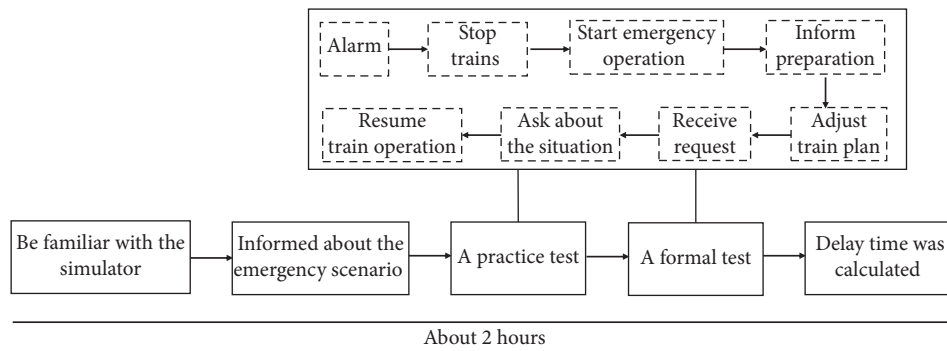


FIGURE 5: The procedures of the simulator test.

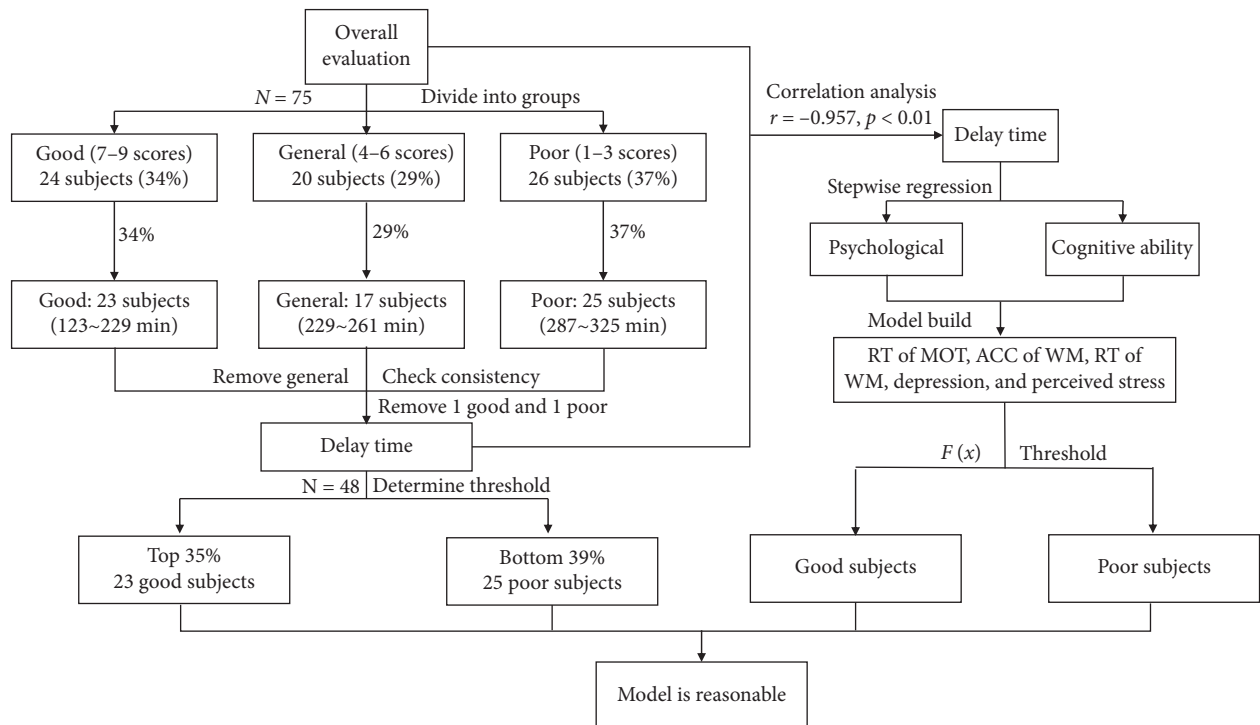


FIGURE 6: Data analysis procedures.

TABLE 2: Descriptive statistics of all variables ($n = 70$).

Scale	Mean	SD	Cronbach's alphas	Category 1 (%)	Category 2 (%)
Age (years)	34.76	3.25	—		
Dispatching experience	4.51	2.22	—		
Education	—	—	—	82	18
Marriage	—	—	—	89	11
ACC of MOT	0.61	0.15	—		
RT of MOT (ms)	1288.04	280.18	—		
ACC of WM	0.67	0.22	—		
RT of WM (ms)	743.39	200.29	—		
Depression	42.75	11.57	0.757		
Anxiety	42.84	11.95	0.82		
Perceived stress	18.7	7.51	0.84		
Social support	55.56	16.91	0.88		
Overall evaluation	5.33	2.31	—		
Delay time (min)	243.31	55.87	—		

were married); (3) independent t -test showed that they did not affect other variables ($p_s > 0.05$).

3.2. Group Division and Threshold Determination. The overall evaluation was a comprehensive evaluation based on the dispatcher's aggregate performance which was susceptible to the personal preferences of the supervisors. Therefore, the overall evaluation had certain subjectivity and arbitrariness. Additionally, the delay times indicated the dispatching ability of the participants to deal with emergency event by operating the simulator, which had certain volatility. As a result, the subjective indicator and the objective indicator were combined into the classification of the dispatchers and the threshold was jointly determined.

Firstly, the overall evaluation was used to stratify the groups. In order to make a large dispersion between groups, the dispatchers were stratified into three groups: Good (7–9 scores), General (general performance) (4–6 scores), and Poor (1–3 scores). In particular, 24 (34%) were Good, 20 (29%) were General, and 26 (37%) were Poor, respectively. The General group was a transition part between the Good and the Poor ones, with the possibility to interfere or to misjudge the results. In this case, the General group was removed from the groups.

The participants were stratified by the overall evaluation into two groups and they were sorted according to the delay time from small to large. Since the subjective indicator had certain subjectivity and arbitrariness, the delay time was used to examine the group consistent with overall evaluation. The top 34% of the delay time was stratified into the Good group, and the bottom 37% was stratified into the Poor group. The participant would be removed when the group of delay time was inconsistent with the group of overall evaluation. In this case, a total of 2 (4%) participants were removed, indicating that the subjective indicator and the objective indicator are of practical unanimity ($r = -0.957$, $p < 0.01$). In the end, the final sample of 48 dispatchers was obtained from 70 valid data, including 23 Good and 25 Poor ones.

Finally, to make clear the predicted values obtained by the stepwise model in Section 3.4 belong to the Good or Poor

group, it was necessary to determine the threshold. Given the groups had been consistently confirmed by overall evaluation and delay time, the maximum delay time (229 min) in the Good was adopted as the threshold. When the predicted value was 229 min or less, the participant was part of the Good group. In other words, when the predicted value was greater than 229 min, the participant was classified into the Poor group.

After grouping, the results of the difference in the demographic variables and the cognitive variables between Good and Poor are presented in Table 3. Two demographic variables, four cognitive variables, and four mental variables were significant differences between Good and Poor ($p_s < 0.05$).

3.3. Correlation of Variables. Two-tailed Pearson correlation analysis was performed to examine the relationship among independent variables (including demographic variables, cognitive variables, and mental variables) and dependent variables (both subjective and objective indicators). As shown in Table 4, the correlation between all independent variables and dependent variables was significant with absolute correlation coefficients ranging from 0.35 to 0.91 ($p_s < 0.05$), which provided basis for establishing regression model for next step. And significant correlations were found between most independent variables. This indicated a multicollinearity relationship between these variables, and this provided basis for stepwise regression. There was a strong correlation between overall evaluation and delay time ($r = -0.95$, $p < 0.01$); therefore, we only used the delay time (objective indicator) as dependent variable in the predictive model.

3.4. Stepwise Regression Model. In this study, the stepwise regression model was established to identify the variables contributing to the delay time. Eleven variables were entered in the process of stepwise and a total of five regression models were established. The model with the best fitting effect which is presented in Table 5 contained five variables: ACC of WM ($\beta = -0.33$, $p < 0.01$), RT of WM ($\beta = 0.27$,

TABLE 3: Comparison of demographic variables and cognitive variables between two groups ($n = 48$).

Measures	Group (mean and standard deviation)		Statistics		P value
	Good (N = 23)	Poor (N = 25)	Good (N = 23)	Poor (N = 25)	
Age	36.35 ± 4.31	33.84 ± 2.01	2.55 ^a	0.02*	
Dispatching experience	5.61 ± 1.99	2.52 ± 0.92	6.80 ^a	0.00**	
ACC of MOT	0.730 ± 0.09	0.53 ± 0.14	5.65 ^a	0.00**	
RT of MOT (ms)	1070.91 ± 183.02	1522.46 ± 237.96	-7.32 ^a	0.00**	
ACC of WM	0.92 ± 0.12	0.47 ± 0.06	15.66 ^a	0.00**	
RT of WM (ms)	595.09 ± 152.0	948.14 ± 32.86	-10.91 ^a	0.00**	
Depression	31.70 ± 6.28	52.20 ± 7.42	-10.29 ^a	0.00**	
Anxiety	34.02 ± 8.30	51.65 ± 10.96	-6.24 ^a	0.00**	
Perceived stress	13.65 ± 5.80	22.04 ± 7.20	-4.42 ^a	0.00**	
Social support	64.22 ± 11.62	52.40 ± 18.20	2.70 ^a	0.01*	

^aT test.

TABLE 4: Correlation between variables.

Variable	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	Y ₁	Y ₂
X ₁	—											
X ₂	0.37*	—										
X ₃	0.43**	0.44**	—									
X ₄	-0.36*	-0.53**	-0.51**	—								
X ₅	0.31*	0.74**	0.58**	-0.72**	—							
X ₆	-0.18	-0.67**	-0.57**	0.59**	-0.78**	—						
X ₇	-0.21	-0.66**	-0.61**	0.66**	-0.84**	0.73**	—					
X ₈	-0.04	-0.34*	-0.40**	0.44**	-0.67**	0.61**	0.77**	—				
X ₉	-0.06	-0.29*	-0.43**	0.50**	-0.56**	0.53**	0.73**	0.70**	—			
X ₁₀	0.10	0.31*	0.28	-0.47**	0.33**	-0.38**	-0.41**	-0.41**	-0.44**	—		
Y ₁	0.37**	0.71**	0.66**	-0.75**	0.90**	-0.85**	-0.81**	-0.64**	-0.53**	0.42**	—	
Y ₂	-0.35*	-0.77**	-0.65**	0.75**	-0.91**	0.84**	0.87**	0.67**	0.53**	-0.36*	-0.95**	—

X₁: age; X₂: dispatching experience; X₃: ACC of MOT; X₄: RT of MOT; X₅: ACC of WM; X₆: RT of WM; X₇: depression; X₈: anxiety; X₉: perceived stress; X₁₀: social support; Y₁: overall evaluation; Y₂: delay time; * $p < 0.05$; ** $p < 0.01$.

TABLE 5: Stepwise regression coefficients (beta weights) for the prediction of delay time.

Model	Constant	β	P value	Adjusted R ²
Constant	128.722		0.00	
ACC of WM	-89.825	-0.33	0.00	
RT of WM	0.085	0.27	0.00	
RT of MOT	0.036	0.17	0.01	0.91
Depression	2.169	0.41	0.00	
Perceived stress	-1.501	-0.18	0.00	
Model	SS	DF	MS	F
Regression	184884.57	5	36976.91	
Residual	16038.41	42	381.87	96.83**
Total	200922.98	47		

* $p < 0.05$; ** $p < 0.01$.

$p < 0.01$), RT of MOT ($\beta = 0.17$, $p < 0.05$), depression ($\beta = 0.41$, $p < 0.01$), and perceived stress ($\beta = -0.18$, $p < 0.01$), accounting for 91% of the total variance of the delay time. The result also showed the quadratic sum of regression and the residual.

In addition, Table 5 revealed that lower ACC of WM ($\beta = -0.33$, $p < 0.01$), longer RT of WM ($\beta = 0.27$, $p < 0.01$), longer RT of MOT ($\beta = 0.17$, $p < 0.05$), higher depression ($\beta = 0.41$, $p < 0.01$), and lower perceived stress ($\beta = -0.18$, $p < 0.01$) were significantly correlated with longer delay time

which indicates worse job performance. According to the stepwise regressions results, we use a statistical model that estimates the relationship between cognitive abilities, mental health, and job performance:

$$Y = 128.722 + 0.036 X_4 - 89.825 X_5 + 0.085 X_6 + 2.169 X_7 - 1.501 X_9, \tag{1}$$

where 128.722 denotes the constant. The vector X₄ is RT of MOT, X₅ is ACC of WM, X₆ is RT of WM, X₇ is depression, and X₉ is perceived stress.

3.5. Model Test. In order to test the model of cognitive abilities, mental factors, and job performance on stepwise regression, five independent variables (ACC of WM, RT of WM, ACC of MOT, depression, and perceived stress) were considered back into the regression model. Therefore, each participant got a new predicted value of delay time. As shown in Figure 7, the threshold value (229 min) and the delay time including the original value and the predicted value were presented. In Figure 7, we observed the predicted value wandered around original value, suggesting that there was no significant difference between them. Moreover, there was a significant jump of value in 229 min between Good and Poor.

In fact, there might be a misclassification of Good and Poor. For example, there were two types of statistical errors:

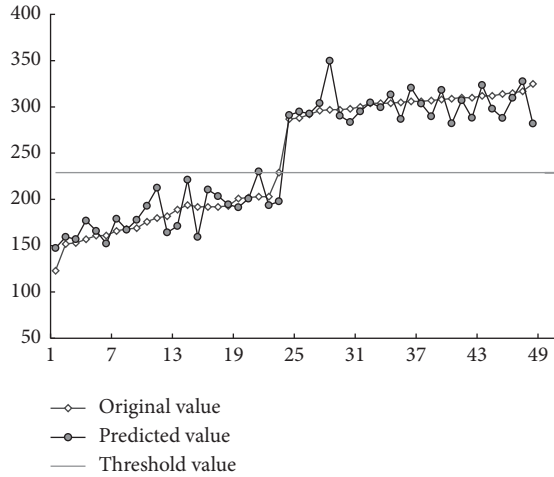


FIGURE 7: Comparison of the difference in delay time between the original and the predicted values.

the RWT errors and type B error, which are common errors. In order to get a more detailed understanding of the accuracy of the stepwise regression model for the division of Good or Poor, the predicted value of the participant was grouped according to the threshold value. As shown in Table 6, 96% of the Good were correctly predicted, while 100% accurate for the Poor.

From Figure 7 and Table 6, it was apparent that the number 21 was misclassified. In order to visually present the results of difference in participant no: 21 and groups between demographic variables, cognitive variables, and mental variables, the radar map was considered. Given the disunity of scale for all variables, the min-max normalization was conducted to solve the problem by the following formula:

$$\widetilde{x}_{ij} = \frac{x_{ij} - \min_i \{x_{ij}\}}{\max_i \{x_{ij}\} - \min_i \{x_{ij}\}}, \quad (2)$$

where $i = 1, 2, 3, \dots, 48, j = 1, 2, \dots, 10$. By this formula, all variables were unified from 0 to 1. Figure 8 presents the results of mean of each variable after unified scale of Good and Poor in full line and the number 21 in imaginary line. The features of number 21 were more consistent with Good. However, compared with the mean of Good, the participant's ACC of WM, RT of WM, RT of MOT, depression, and perceived stress were obviously different. The five indicators were used to predict the job performance. Therefore, the participant was misclassified.

4. Discussion

The goal of the current study is to examine the relationship between the cognitive abilities, mental factors, and the job performance of the Chinese HSR dispatchers. Findings showed that the multiobject ability, the working memory ability, and 2 mental factors have strong relations with the job performance (indicated by delay time), and tremendous distinctions between groups of good and poor performance of the HSR dispatchers were revealed. In addition, the

TABLE 6: Prediction rate of regression model between the two groups.

Predicted type Original type	Good (%)	Poor (%)
Good (%)	96	4
Poor (%)	0	100

stepwise regression modelling is employed to construct the regression equation, in which the demographic data, the accuracy and the reaction time of working memory, the reaction time of multiobject ability, the depression, and the perceived stress appeared as the independent variables and the job performance appeared as the dependent variable. The current research findings can offer indicators and models in future HSR dispatchers' selection.

Consistent with existing studies [24, 69], the positive relationship between cognitive abilities and dispatchers' job performance is supported in the current study. Working memory, which included both short-term storage and central executive control [29], is exactly what the dispatcher's job requires according to task analysis (Table 1). It is expected and understandable that working memory can significantly predict HSR dispatcher's key performance, just as working memory has demonstrated to be a good predictor for shooting behavior [70], school grades [71], and other performance [72]. HSR dispatchers should also dynamically monitor the operational status of multiple trains as well as the conditions of signals, stations, switchers, and weather, which involves dynamic visual tracking of multiobjects [73]. That leads to the ability of multiobject tracking, becoming an important predictor of job performance just as the similar job of air traffic controllers [30].

The result that mental factors were also predictors of the job performance of HSR dispatchers in this study is expected, which is consistent with results of other related occupations in previous studies [33, 34, 74]. Depression, as the most concerned mental health symptom, is associated with low self-esteem, reduced work motivation, and slower cognitive processing [33, 75]. For HSR dispatchers, depression could directly affect their subjective and objective performance indicators on the one hand. On the other hand, it could damage cognitive abilities including both WM and MOT, indirectly affecting the job performance. Although overload pressure may always accompany with fatigue, insomnia, neurological disorders, and other symptoms that cause worse performance [76], moderate pressure could also promote work motivation which leads to even better performance [74]. In the current study, perceived stress was measured. The results of more perceived stress are associated with less delay time which means better job performance. It was most probably because of the particularities of HSR dispatching work that need continuous vigilance or the relatively moderate perceived stress score (mean = 18.02 and SD = 7.76). Inconsistent to other studies, anxiety and perceived social support did not enter into the predictive model of job performance [77, 78]. This could be understandable from aspect of job analysis. In actual work, there are not many cases which require emergency dispatching though

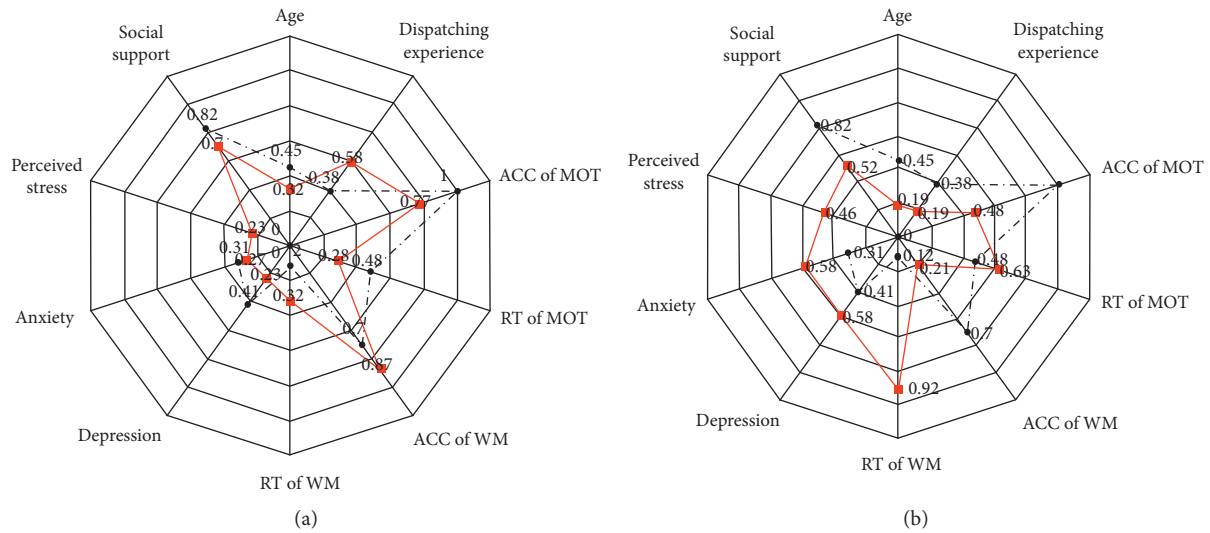


FIGURE 8: The results of mean of each variable of Good and Poor and the number 21.

constant vigilance is needed, while anxiety is mainly related to emergency situations or unmanageable events. Besides, the perceived social support scores obtained in this study were relatively concentrated in high section, while the ceiling effect makes this variable could not distinguish between the Good and Poor groups.

To summarize, the current study establishes a prediction model for the job performance of Chinese HSR dispatchers. The quantitative relationship between cognitive abilities, mental factors, and the job performance, to the best knowledge of the authors, has rarely been examined in previous study. It confirms that cognitive abilities including multiobject tracking and working memory and mental factors are important for the job performance among the existing qualified HSR dispatchers. The predictive variables including the ACC of WM, the RT of WM, the ACC of MOT, depression, and perceived stress explained an additional of 91% of the variance in the job performance. Moreover, the job performance model gets the accuracy up to 96%. These findings suggest that tests of the 2 cognitive abilities and 2 mental factors may be useful tools complementary to the current practice of HSR dispatcher selection and assessment.

Although the traditional selection and assessment can investigate the dispatcher's dispatching knowledge and professional skills to a certain extent, the operational adaptability and mental factors might be more accurate potential indicators which have not yet been considered in terms of affecting the dispatcher's operational behavior safety especially in emergency scenario. Taking an emergency scenario of the dispatcher encountered an abnormal train noise, for example. However, in order to maximize the train operation, dispatchers should perform multiple tasks and organize railway personnel with different responsibilities quickly (even in seconds) and accurately. This makes dispatchers' basic cognitive abilities mainly working memory and multiobject tracking become the key factor of impacting their job performance. Regarding mental factors,

the state of depression might lead to slower or less accurate reaction that would produce irreversible consequence, while perceived moderate stress would help to maintain alertness and promote more validity operation. Only when more comprehensive impact factors are included in actual HSR dispatcher's selection and assessment, our daily transportation will be safer. Therefore, building the prediction model taking both cognitive and mental factors into account is not only of theoretical significance but also of practical significance.

A limitation of the current study is that the HSR dispatchers were all from one group, China Railway Chengdu Group. Future studies are needed to examine the predictive model among other dispatcher populations such as other geographical regions in China Railway and other countries. In addition to mental factors such as depression and perceived stress, other psychological factors such as personality and motivation may also be good predictors of job performance [79]. Future studies are needed to examine these additional factors and their joint influence on the HSR dispatchers' job performance.

5. Conclusions

In conclusion, both cognitive factors specifically multiobject visual tracking and working memory and mental factors such as depression and perceived stress are significant predictors of the HSR dispatchers' job performance. The results from the current study provide a theoretical basis for the training of HSR dispatchers and can support future work towards establishing standardized methods and criteria for dispatcher selection and management. Additionally, the subjective and objective indicators have a high degree of consistency, which also provides a theoretical basis for the selection tools. In the next step, we will try other methods (data mining, artificial intelligence, etc.) to explore the relationship among these cognitive abilities and much more psychological factors.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the Science and Technology Program of Sichuan Province (2019YFG0043), Science and Technology Program of China Railway (2018F024), National Natural Science Foundation of China (51108390 and 71601163), and 100 Talents Program Project of Southwest Jiaotong University (grant no. 2682017WBR18).

References

- [1] H. Huang, J. Chen, F. Yu, and Z. Zhu, "Establishing the enterprises' innovation ecosystem based on dynamics Core competence-the case of China's high-speed railway," *Emerging Markets Finance and Trade*, vol. 55, no. 4, pp. 843–862, 2019.
- [2] X. Dong, "High-speed railway and urban sectoral employment in China," *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 603–621, 2018.
- [3] Y. Fan, Z. Li, J. Pei, H. Li, and J. Sun, "Applying systems thinking approach to accident analysis in China: case study of "7.23" Yong-Tai-Wen High-Speed train accident," *Safety Science*, vol. 76, pp. 190–201, 2015.
- [4] W. Wang, X. Liu, and Y. Qin, "A modified HEART method with FANP for human error assessment in high-speed railway dispatching tasks," *International Journal of Industrial Ergonomics*, vol. 67, pp. 242–258, 2018.
- [5] W. U. Hai-Tao and X. Luo, "Research on human error probability quantification method for high-speed railway train dispatcher," *China Safety Science Journal*, vol. 67, 2015.
- [6] F. Chu, Y. Fu, and S. Liu, "Organization is also a "life form": organizational-level personality, job satisfaction, and safety performance of high-speed rail operators," *Accident Analysis & Prevention*, vol. 125, pp. 217–223, 2019.
- [7] S. Kalkavan and A. Katrinli, "The effects of managerial coaching behaviors on the employees' perception of job satisfaction, organisational commitment, and job performance: case study on insurance industry in Turkey," *Procedia - Social and Behavioral Sciences*, vol. 150, pp. 1137–1147, 2014.
- [8] F. P. Morgeson, K. Delaney-Klinger, and M. A. Hemingway, "The importance of job autonomy, cognitive ability, and job-related skill for predicting role breadth and job performance," *Journal of Applied Psychology*, vol. 90, no. 2, pp. 399–406, 2005.
- [9] Y. Yao, R. Wang, and K. Y. Wang, "The influence of emotional intelligence on job performance: moderating effects of leadership," in *Proceedings of 2009 International Conference on Management Science and Engineering*, vol. 14-16, pp. 1155–1160, New York, NY, USA, 2009.
- [10] J. D. Pincus, "Communication satisfaction, job satisfaction, and job performance," *Human Communication Research*, vol. 12, no. 3, pp. 395–419, 1986.
- [11] G. Dokko, S. L. Wilk, and N. P. Rothbard, "Unpacking prior experience: how career history affects job performance," *Organization Science*, vol. 20, no. 1, pp. 51–68, 2009.
- [12] T. R. Carretta, "Pilot candidate selection method. aviation psychology and applied human factors," 2011.
- [13] D. Gardner and D. L. Deadrick, "Underprediction of performance for US minorities using cognitive ability measures," *Equal Opportunities International*, vol. 27, no. 5, pp. 455–464, 2008.
- [14] A. R. Kolz, L. A. McFarland, and S. B. Silverman, "Cognitive ability and job experience as predictors of work performance," *The Journal of Psychology*, vol. 132, no. 5, pp. 539–548, 1998.
- [15] K. Anstey, J. Wood, S. Lord, and J. Walker, "Cognitive, sensory and physical factors enabling driving safety in older adults," *Clinical Psychology Review*, vol. 25, no. 1, pp. 45–65, 2005.
- [16] A. Blane, H. Lee, T. Falkmer, and T. Dukic Willstrand, "Cognitive ability as a predictor of task demand and self-rated driving performance in post-stroke drivers - implications for self-regulation," *Journal of Transport & Health*, vol. 9, pp. 169–179, 2018.
- [17] D. B. Devoe, "An analysis of the job of railroad train dispatcher," 1974.
- [18] J. Gertler, "Selection of railroad dispatcher candidates," 2003.
- [19] E. M. Roth, N. Malsch, and J. Multer, "Understanding how train dispatchers manage and control trains: results of a cognitive task analysis," 2001.
- [20] W. C. Periman, "Relationship of working memory to job performance and innovation with stress and effort as moderators," 2016.
- [21] R. W. Engle, "Working memory capacity as executive attention," *Current Directions in Psychological Science*, vol. 11, no. 1, pp. 19–23, 2002.
- [22] G. Salvatore, B. R. Cornwell, F. Sambataro et al., "Anterior cingulate desynchronization and functional connectivity with the amygdala during a working memory task predict rapid antidepressant response to ketamine," *Neuropsychopharmacology*, vol. 35, no. 7, pp. 1415–1422, 2010.
- [23] A. M. Owen, J. Doyon, M. Petrides, and A. C. Evans, "Planning and spatial working memory: a positron emission tomography study in humans," *European Journal of Neuroscience*, vol. 8, no. 2, pp. 353–364, 1996.
- [24] J. Wei, T. Feng, and P. Jing, "Features of multiple-objects tracking task processing in student pilots," *Space Medicine & Medical Engineering*, vol. 35, 2013.
- [25] R. Colom, A. Martínez-Molina, P. C. Shih, and J. Santacreu, "Intelligence, working memory, and multitasking performance," *Intelligence*, vol. 38, no. 6, pp. 543–551, 2010.
- [26] A. M. Owen, K. M. McMillan, A. R. Laird, and E. Bullmore, "N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies," *Human Brain Mapping*, vol. 25, no. 1, pp. 46–59, 2005.
- [27] S. Cao, U. Reiter, and M. Weitzel, "Influence of interaction on perceived quality in audio visual applications: subjective assessment with n-back working memory task," in *Proceedings of Audio Engineering Society Conference: 30th International Conference: Intelligent Audio Environments*, New York, NY, USA, 2010.
- [28] R. Allen, P. McGeorge, D. Pearson, and A. B. Milne, "Attention and expertise in multiple target tracking," *Applied Cognitive Psychology*, vol. 18, no. 3, pp. 337–347, 2004.
- [29] A. Baddeley, "Working memory: looking back and looking forward," *Nature Reviews Neuroscience*, vol. 4, no. 10, pp. 829–839, 2003.

- [30] R. M. Hope, E. M. Rantanen, and L. Oksama, "Multiple identity tracking and entropy in an ATC-like task," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 54, no. 13, pp. 1012–1016, 2010.
- [31] S. Harenberg, R. McCaffrey, M. Butz et al., "Can multiple object tracking predict laparoscopic surgical skills?" *Journal of Surgical Education*, vol. 73, no. 3, pp. 386–390, 2016.
- [32] I. Legault, R. Allard, and J. Faubert, "Healthy older observers show equivalent perceptual-cognitive training benefits to young adults for multiple object tracking," *Frontiers in Psychology*, vol. 4, 2013.
- [33] M. F. Hilton, Z. Staddon, J. Sheridan, and H. A. Whiteford, "The impact of mental health symptoms on heavy goods vehicle drivers' performance," *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 453–461, 2009.
- [34] C. M. Wickens, R. G. Smart, and R. E. Mann, "The impact of depression on driver performance," *International Journal of Mental Health and Addiction*, vol. 12, no. 4, pp. 524–537, 2014.
- [35] M. L. Cunningham and M. A. Regan, "Are happy drivers better drivers?" *The Impact of Emotion, Life Stress and Mental Health Issues on Driving Performance and Safety*, vol. 12, 2017.
- [36] K. Suzuki, T. Ohida, Y. Kaneita et al., "Mental health status, shift work, and occupational accidents among hospital nurses in Japan," *Journal of Occupational Health*, vol. 46, no. 6, pp. 448–454, 2004.
- [37] J. A. Edwards, A. Guppy, and T. Cockerton, "A longitudinal study exploring the relationships between occupational stressors, non-work stressors, and work performance," *Work & Stress*, vol. 21, no. 2, pp. 99–116, 2007.
- [38] M. Bubonya, D. A. Cobb-Clark, and M. Wooden, "Mental health and productivity at work: does what you do matter?" *Labour Economics*, vol. 12, 2016.
- [39] S. Mulder and D. de Rooy, "Pilot mental health, negative life events, and improving safety with peer support and a just culture," *Aerospace Medicine and Human Performance*, vol. 89, no. 1, pp. 41–51, 2018.
- [40] A. Hatami, S. Vosoughi, A. F. Hosseini, and H. Ebrahimi, "Effect of Co-driver on job content and depression of truck drivers," *Safety and Health at Work*, vol. 10, no. 1, pp. 75–79, 2019.
- [41] T. A. Brown, B. F. Chorpita, W. Korotitsch, and D. H. Barlow, "Psychometric properties of the depression anxiety stress scales (DASS) in clinical samples," *Behaviour Research and Therapy*, vol. 35, no. 1, pp. 79–89, 1997.
- [42] L. Liu and X. Dong, "Study on Mental Health status and influencing factors of OCC dispatchers in S City," *Human Resource Management*, vol. 2, 2017.
- [43] T. A. Raslear, "Preliminary examination of railroad dispatcher workload, stress, and fatigue," 2002.
- [44] N. Bolger, A. DeLongis, R. C. Kessler, and E. A. Schilling, "Effects of daily stress on negative mood," *Journal of Personality and Social Psychology*, vol. 57, no. 5, pp. 808–818, 1989.
- [45] C. Sandi, "Stress and cognition," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 4, no. 3, pp. 245–261, 2013.
- [46] O.-I. Siu, C.-q. Lu, and P. E. Spector, "Direct and indirect relationship between social stressors and job performance in Greater China: the role of strain and social support," *European Journal of Work and Organizational Psychology*, vol. 22, no. 5, pp. 520–531, 2013.
- [47] M.-L. Wang and L.-J. Tsai, "Work-family conflict and job performance in nurses," *Journal of Nursing Research*, vol. 22, no. 3, pp. 200–207, 2014.
- [48] A. A. Beaujean, S. Parker, and X. Qiu, "The relationship between cognitive ability and depression: a longitudinal data analysis," *Social Psychiatry and Psychiatric Epidemiology*, vol. 48, no. 12, pp. 1983–1992, 2013.
- [49] C. G. La Fleur and T. A. Salthouse, "Which aspects of social support are associated with which cognitive abilities for which people?" *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 72, pp. 1006–1016, 2017.
- [50] B. Coy, W. H. O'Brien, T. Tabaczynski, J. Northern, and R. Carels, "Associations between evaluation anxiety, cognitive interference and performance on working memory tasks," *Applied Cognitive Psychology*, vol. 25, no. 5, pp. 823–832, 2011.
- [51] T. Ortiz, J. Perez-Serrano, C. Zaglul et al., "Deficit of cognitive event-related potentials during a working task in patients with major depression," *Actas Esp Psiquiatr*, vol. 31, pp. 177–181, 2003.
- [52] C.-Y. Chen, J. P. Lawlor, A. K. Duggan, J. B. Hardy, and W. W. Eaton, "Mild cognitive impairment in early life and mental health problems in adulthood," *American Journal of Public Health*, vol. 96, no. 10, pp. 1772–1778, 2006.
- [53] J. A. Hadwin and H. J. Richards, "Working memory training and CBT reduces anxiety symptoms and attentional biases to threat: a preliminary study," *Frontiers in Psychology*, vol. 7, 2016.
- [54] S. L. Hatch, P. B. Jones, D. Kuh, R. Hardy, M. E. J. Wadsworth, and M. Richards, "Childhood cognitive ability and adult mental health in the British 1946 birth cohort," *Social Science & Medicine*, vol. 64, no. 11, pp. 2285–2296, 2007.
- [55] L. T. Martin, L. D. Kubzansky, K. Z. LeWinn, L. P. Lipsitt, P. Satz, and S. L. Buka, "Childhood cognitive performance and risk of generalized anxiety disorder," *International Journal of Epidemiology*, vol. 36, no. 4, pp. 769–775, 2007.
- [56] S. Zammit, P. Allebeck, A. S. David et al., "A longitudinal study of premorbid IQ score and risk of developing Schizophrenia, Bipolar disorder, severe depression, and other nonaffective psychoses," *Archives of General Psychiatry*, vol. 61, no. 4, pp. 354–360, 2004.
- [57] D. R. Johnson and S. D. Gronlund, "Individuals lower in working memory capacity are particularly vulnerable to anxiety's disruptive effect on performance," *Anxiety, Stress & Coping*, vol. 22, no. 2, pp. 201–213, 2009.
- [58] K. Smith, D. Gatica-Perez, J. Odobez, and B. Sileye, "Evaluating multi-object tracking," in *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, vol. 21–23, p. 36, New York, NY, USA, 2005.
- [59] G. A. Alvarez, T. S. Horowitz, H. C. Arsenio, J. S. DiMase, and J. M. Wolfe, "Do multielement visual tracking and visual search draw continuously on the same visual attention resources?" *Journal of Experimental Psychology: Human Perception and Performance*, vol. 31, no. 4, pp. 643–667, 2005.
- [60] W. W. K. Zung, C. B. Richards, and M. J. Short, "Self-rating depression scale in an outpatient clinic," *Archives of General Psychiatry*, vol. 13, no. 6, pp. 508–515, 1965.
- [61] D. A. Dunstan, N. Scott, and A. K. Todd, "Screening for anxiety and depression: reassessing the utility of the Zung scales," *Bmc Psychiatry*, vol. 17, p. 329, 2017.
- [62] W. ., R. LINDSAY and A. ., M. MICHIE, "Adaptation of the Zung self-rating anxiety scale for people with a mental handicap," *Journal of Mental Deficiency Research*, vol. 32, pp. 485–490, 2010.
- [63] S. Cohen, T. Kamarck, and R. Mermelstein, "A global measure of perceived stress," *Journal of Health and Social Behavior*, vol. 24, no. 4, pp. 385–396, 1983.

- [64] X. D. Wang, X. L. Wang, and H. Ma, "Manual of mental health rating scale," *Chinese Mental Health Journal*, vol. 13, pp. 31–35, 1999.
- [65] G. D. Zimet, N. W. Dahlem, S. G. Zimet, and G. K. Farley, "The multidimensional scale of perceived social support," *Journal of Personality Assessment*, vol. 52, no. 1, pp. 30–41, 1988.
- [66] J. P. Meyer, S. V. Paunonen, I. R. Gellatly, R. D. Goffin, and D. N. Jackson, "Organizational commitment and job performance: it's the nature of the commitment that counts," *Journal of Applied Psychology*, vol. 74, no. 1, pp. 152–156, 1989.
- [67] T. A. Wright, R. Cropanzano, and D. G. Meyer, "State and trait correlates of job performance: a tale of two perspectives," *Journal of Business and Psychology*, vol. 18, pp. 365–383, 2004.
- [68] W. Darr and G. Johns, "Political decision-making climates: theoretical processes and multi-level antecedents," *Human Relations*, vol. 57, no. 2, pp. 169–200, 2004.
- [69] K. Barker, R. Allen, and P. McGeorge, "Multiple-object tracking," *Experimental Psychology*, vol. 57, no. 3, pp. 208–214, 2010.
- [70] G. A. Brewer, B. Hunter Ball, and J. M. Ware, "Individual differences in working memory capacity and shooting behavior," *Journal of Applied Research in Memory and Cognition*, vol. 5, no. 2, pp. 185–191, 2016.
- [71] S. Krumm, M. Ziegler, and M. Buehner, "Reasoning and working memory as predictors of school grades," *Learning and Individual Differences*, vol. 18, no. 2, pp. 248–257, 2008.
- [72] A. Mojzisch, S. Krumm, and T. Schultze, "Do high working memory groups perform better?" *Journal of Personnel Psychology*, vol. 13, no. 3, pp. 134–145, 2014.
- [73] A. Tazoniero, R. Gonçalves, and F. Gomide, "Decision making strategies for real-time train dispatch and control," in *Analysis and Design of Intelligent Systems Using Soft Computing Techniques*, P. Melin, O. Castillo, E. G. Ramirez, J. Kacprzyk, and W. Pedrycz, Eds., Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 195–204, 2007.
- [74] G. Mathur, S. Vigg, S. Sandhar, and U. Holani, "Stress as a correlate of job performance: a study of manufacturing organizations," *Journal of Advances in Management Research*, vol. 4, no. 2, pp. 79–85, 2007.
- [75] T. A. Wright and D. G. Bonett, "Growth coping, work satisfaction and turnover: a longitudinal study," *Journal of Business and Psychology*, vol. 6, no. 1, pp. 133–145, 1991.
- [76] N. Abdoli, V. Farnia, A. Delavar et al., "Mental health status, aggression, and poor driving distinguish traffic offenders from non-offenders but health status predicts driving behavior in both groups," *Neuropsychiatric Disease and Treatment*, vol. 11, pp. 2063–2070, 2015.
- [77] P. Branscum, T. Haider, D. Brown, and M. Sharma, "Using emotional intelligence and social support to predict job performance of health educators," *American Journal of Health Education*, vol. 47, no. 5, pp. 309–314, 2016.
- [78] T. G. Reio and J. L. Callahan, "Affect, curiosity, and socialization-related learning: a path analysis of antecedents to job performance," *Journal of Business and Psychology*, vol. 19, pp. 3–22, 2004.
- [79] M. Guo, W. Wei, G. Liao, and F. Chu, "The impact of personality on driving safety among Chinese high-speed railway drivers," *Accident Analysis & Prevention*, vol. 92, pp. 9–14, 2016.