Spatiotemporal Analysis of Competition Between Subways and Taxis Based on Multi-Source Data

RUI WANG[1], FENG CHEN[1], XIAOBING LIU[2], AND TAKU FUJIYAMA[3]
[1]School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China
[2]MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[3]Department of Civil, Environmental and Geomatic Engineering, University College London, London WC1E 6BT, U.K.

Corresponding author: Feng Chen (fengchen@bjtu.edu.cn)

ABSTRACT Excessive competition between taxis and subways has eroded the advantages of public transit systems such as worsening road traffic congestion and environment. This study aims to improve the appeal of subways by a comprehensive understating of competition between taxis and subways. We investigate competitive relationship between these two transportation modes by using empirical multi-source data. First, non-negative matrix factorization (NMF) algorithm is used to discover the spatiotemporal travel patterns of subway-competing taxi users (SCTUs). Second, we propose a new index to quantify the competitiveness of subways based on the actual mode choices results. Then, we reveal the spatiotemporal heterogeneity of competitiveness from perspective of subway network. Taking Beijing, China, for a case study, we extract a week's worth of GPS records on taxi trajectory and smartcard data of subways. Subway-competing taxi trips (SCTTs) account for the largest proportion of the total taxi trips. As a result, three basic patterns are found in SCTTs. Subway station pairs with high and less competition are divided according to competitiveness index. Among low competition station pairs, three spatial structures are observed, including low-competition collinearity corridors, radial communities, and links between paralleled subway lines. Combining the distribution results of travel pattern and competitiveness degree, short-term and long-term planning suggestions are recommended respectively for station pairs with high demand but low competitiveness and those with low demand and low competitiveness. These findings provide useful insights into promoting more effective and sensitive policies to balance the competition and attract more taxi passengers to the subway system.

INDEX TERMS Subway-competing taxi, sustainable transportation, non-negative matrix factorization, competition, subway planning.

I. INTRODUCTION

Road congestion and automobile pollution are two serious challenges in cities worldwide. To mitigate these problems, countries have taken various measures including promoting sustainable public transport, which is one of the most effective ways at present [10], [13], [23]. In urban areas, due to the dense population and huge travel demand, various public transit modes are operated simultaneously, such as subway and taxi. Generally, subway is the dominant mode of the metropolitan public transit system because of its advantages of high capacity, low carbon footprint, and punctuality. By contrast, taxis barely offer all those advantages of public transport due to their lower capacity and considerable

The associate editor coordinating the review of this manuscript and approving it for publication was Chien-Ming Chen.

deadhead kilometers while locating passengers in busy business areas. They help neither relieve road congestion nor reduce air pollution [1], [13]. While, the optimal purpose of a taxi is to provide a feeder service to subways or buses, and serve as a substitute where private driving or other public transit modes are not feasible [3], [40]. Therefore, they still play an indispensable role in urban transportation. However, due to the relative inconvenience of subways or because passengers seek travel comfort, taxis are increasingly wresting the share of passenger flow from subways rather than serving as a feeder to the latter. In fact, taxis are giving excessive competition to subways; this runs counter to the optimal purpose of taxi services and erodes the promotion of sustainable transport. To balance the market share of subway and taxi, factors which influence mode choices are recognized from the perspective of individual [5], [33]. Incidentally, there are

few studies that have conducted in-depth evaluation of this competition from the perspective of network.

The shortage of studies on interactional relationships of taxis and subways, especially the competition between the two, can be attributed to two reasons. The first is the lack of actual trip data. Most studies on taxi or subway are based on costly survey, which can neither draw the whole picture of the competitive relationship nor identify the specific travel patterns of the associated passengers. With the GPS and the auto fare collection system (AFC), the real trip information of each taxi and subway passengers in urban areas can be recorded, which act as a powerful tool to fill such research gaps as trip-specific analysis of spatiotemporal mobility and exploration of the interplay between taxis and subways [9], [45]. The second reason is the small share of taxi riders relative to that of the public transit system passengers in urban travel. The taxi is regarded as a luxury transit tool in some cities [5]. In Chinese cities, however, it can be afforded by the average person where taxi trips account for a considerable share of the public transit system [24]. Moreover, online transportation services similar to the taxi, such as ride-hailing (*e.g.*, Uber and DiDi), have become popular and their market share is expanding rapidly [26], [27], [31], [32]. Research on the competitive relationship between taxis and subways can encourage ride-hailing users, and even private car owners, to take the subway and balance the competition. Therefore, it is more significant to analyze the competition between taxis and subways in the face of the changing market.

To fill the research gap, this research focuses on the competition between the taxi and the subway from two aspects—travel pattern of competing trips and spatiotemporal analysis of competitiveness in the level of station pairs. It aims at reducing competition between the two travel modes, achieving the optimal operational purpose of taxis, and guaranteeing the share ratio of the subway system relative to the taxi system. The two aspects of competition analysis are prerequisites to increasing the ridership of subways through effective and targeted measures. Using the actual trip data from a taxi's GPS and the subway smartcard in Beijing, China, we conducted a travel pattern analysis of taxi users who are likely to travel by subway and proposed a new competitiveness index to quantify the degree of attractiveness of subway to taxi users in urban areas. The competitiveness was analyzed in the level of station pair. To the best of the authors' knowledge, this study is the first to quantify competition between taxis and subways based on massive data and investigate spatiotemporal heterogeneity of competitiveness in station pair level. The associated results can provide valuable suggestions to mitigate the excess competition in practice.

Compared to the existing research, this study will make significant contributions in the following aspects. (1) Latent basic travel patterns of subway-competing taxi users are identified by NMF for the first time. SCTUs are part of taxi users but they have different basic collective patterns. (2) A new technique of evaluating competitiveness is proposed based on taxi and subway data. Competitiveness index is calculated according to the actual mode choices which is obtained by massive data from taxi GPS and smart cards. The indicator is more accurate to reflect the fact of subway competition rate. (3) The dynamic and heterogeneous features of competitiveness are determined by taking subway station pairs as the unit. By using the proposed new index in this study, spatiotemporal heterogeneity of subway competitiveness is discovered in level of station pairs. (4) By taking Beijing as the case city, we put forward some short-term strategies and long-term subway planning suggestions based on the travel pattern of SCTUs and characteristics of competitiveness.

The rest of the paper is structured as follows. Section 2 presents a literature review about the competitive relationship between subways and taxis and the associated measurement methods. Section 3 describes the datasets from multiple sources, methods of identifying travel patterns, and a model of the competitiveness index. The results of the travel patterns of competing trips and the distribution of competitiveness are analyzed in Section 4. Finally, Section 5 discusses the policy implications of the results and summarizes the main findings of this study.

II. LITERATURE REVIEW
A. RELATIONSHIP BETWEEN SUBWAYS AND TAXIS
Taxi and subway are two important transit modes in urban areas; the interplay between them was explored in previous works, to improve traffic performance and promote sustainable transport. Li *et al.* [22], for example, validated this type of interaction and indicated that a new subway line could influence the spatial distribution of taxi ridership in Wuxi, China. They found a decrease in the volume of taxi trips that could be replaced by a new subway line, while finding an increase in the volume of trips connecting subway stations. Kim [14] showed that subways and taxis served residents synergistically at different scales of the functional structure of Seoul, South Korea. Travel patterns of subway passengers dominated large residential and business districts, whereas taxi users dominated the smaller ones. Yue *et al.* [46] argued that different mass transit modes cooperated or competed based on the demographic and socioeconomic attributes of the underlying urban environments.

To understand this interplay explicitly, many researchers investigated the potential relationships between taxis and public transit based on actual trip data. Austin and Zegras [1] indicated that taxis acted both as a substitute and as a complement to mass transit. Hochmair [11] found that taxis competed with buses in well-served areas as the number of bus stops was negatively correlated to the number of taxi trips. Wang *et al.* [36] defined three relationships based on the spatial relation of taxi trips and fixed-route transit stations through the taxi data of New York, US. Taxi trips with both origin and destination in the catchment area of transit stations are transit-competing trips, and the authors found transit-competing trips occupied the highest proportion of trips, i.e., up to 58.54%. A similar definition was

adopted by Jiang *et al.* [12], who analyzed the relationship between subways and taxis in Beijing and demonstrated that the travel characteristics of trips were different in the three clusters of subway-competing, subway-extending and subway-complement. To the best of the authors' knowledge, the relationship between taxis and transit is broadly classified into three types: extending, competing, complementing. Complementing and extending subway mobility with taxis will positively shift passengers from cars to public transit. However, excessive competition from the taxi to the subway can erode the latter's ridership, which should be refrained from in practice. Therefore, this study focuses on the competitive relationship between subways and taxis, and thus addresses the obstacle to promoting sustainable transport. We assumed that a competitive relation exists where taxi trips can be substituted by a physical subway network.

B. DEFINITION OF COMPETITIVENESS

In previous works, many researchers have aimed to find out an effective way to improve the competitiveness of public transit. Those studies mostly used field surveys on mode choice. They defined competitiveness by the choice among multiple travel modes. Li *et al.* [20], [21] analyzed the main factors influencing the intention of car users to take public transport, and found that comfort, reliability, and economy were the most helpful in encouraging car users to switch their travel mode and enhancing public transit competitiveness. Zou *et al.* [48] proved that family or Policy can guide people who do not own private cars to avail public transportation. For car users, improving the service level of public transit, including comfort, safety, and convenience, is the primary strategy to enhance competitiveness. Martín *et al.* [28] calculated the competition between a high-speed train (HST) and an aircraft by the ratio of the probability of choosing a plane or a train based on survey data. A ratio of 1 or in the small neighborhood of 1 means the aircraft and the HST are equally competitive. A ratio of less than 1 means that the HST is more competitive. The competitiveness values in the above studies were calculated based on the mode choices from stated preference(SP) survey data; these values may be different from the actual behavior on the account of the responses were hypothetical [3].

To measure the competition between different transit modes, studies have proposed quantitative indices based on empirical data. Specifically, travel time was most frequently used as an index to reflect competitiveness. According to these studies, for a certain trip, the transit mode that requires a shorter time was more competitive. Ellison [8] analyzed cycling competitiveness by calculating the trip time. Sun *et al.* [34] compared door-to-door travel time of various transit modes to estimate the competitiveness among them. Lee *et al.* [18] quantified competition according to the proportion of origin–destination (OD) pairs and travel time in a traffic analysis zone. Besides, a few other measures were also used to compute competitiveness. Witlox [39]

evaluated competition between bicycle and automobile from the distributions of trip length and speed in Ghent, Belgium. They found that cycling could only be competitive for trips up to 5 km (or maximum 7.5 km). Casello [2] computed competitiveness as the ratio of generalized cost of traveling by transit to the generalized cost of traveling by automobile for one trip.

In short, the competitiveness index in previous studies was designed frequently from the perspective of travel cost. Regardless of the approach used, i.e., index of travel time or the probability of mode choice, the studies hypothesized that passengers are rational and that they will choose the mode with the lower cost; thus, the mode selected by more passengers was regarded to have higher competitiveness. However, estimating the actual travel cost is hard because of the developed recognition of massive factors, from physical conditions (travel time, distance, and fare) to psychological attitudes (comfort, safety, and environment) [19], [27]. In addition, the weight of each factor may vary with time or passenger [30], [41]–[43]. Consequently, using the travel cost to directly measure competitiveness would introduce deviations. To avoid these problems, we proposed the index of competitiveness based on actual observed mode choices by using the raw data of subway and taxi competing trips.

C. SPATIOTEMPORAL ANALYSIS ON COMPETITION

A deep understanding of the spatiotemporal properties of competition between subways and taxis is necessary to formulate elaborate policies to improve subway competitiveness and encourage more taxi users to take the subway. However, research in this field is scant and still in its nascent stages. Wang and Jiang [12], [36] studied the basic characteristics (travel distance and travel time) of subway-competing taxi trips (SCTTs) and recognized the spatial distribution of pick-up points at the regional level. The travel pattern of SCTTs has not been fully understood in the temporal or spatial dimensions. Moreover, analysis at the fine-grained level should also be explored. Recently, for a high-dimensional taxi dataset, matrix decomposition methods were used to exploit the latent spatiotemporal travel pattern, such as principal components analysis (PCA), singular value decomposition (SVD), and non-negative matrix factorization (NMF) [6], [7], [16]. Especially, NMF is more popular owing to its non-negative character, which is in accordance with the property of travel demand [25], [35].

In addition to studies on the travel pattern of SCTTs, fewer research works have focused on the competitiveness between the two transit modes. Ye *et al.* [44] proposed an index of the tendency toward subways and taxis based on ridership and tried to identify the correlation between the index and factors such as traffic condition, population, and built environment. They understood competition from macroscopic aspects and only concentrated on entire market shares in the city. They did not consider the spatiotemporal heterogeneity of competition.

III. DATA AND METHODS

A. STUDY AREA AND DATASETS

The study area is the inner city within the sixth ring road of Beijing, China (Fig. 1). A megacity with a complex road network, Beijing can be divided into the city center and suburban areas by ring roads. The city center usually refers to the area within the sixth ring road. Most residents in Beijing, i.e., over 21 million, live and work within this area, causing huge public traffic demand. It can be assumed that the vitality of this region represents the whole city of Beijing. However, traffic congestion is a serious challenge in the inner city. To satisfy citizens' travel demands and relieve the traffic pressure, the subway, which is the dominant mode of public transport in Beijing, has been rapidly developing outward from the city center and has already extended to the sixth ring road before this study. By October 2014, there were 17 subway lines (including one airport express) and 233 stations in operation in Beijing. The city center area with both major taxi services and an entire subway network is the most appropriate to explore the competition between subways and taxis.
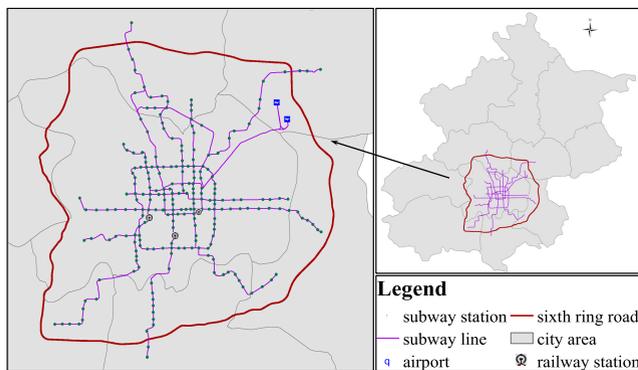

FIGURE 1. Study area and subway layout.

To identify the competition between subways and taxis, three datasets were prepared, i.e., subway network data, smartcard data, and taxi GPS data. Network data including locations of stations and layout of subway lines were obtained from a GIS dataset. Smartcard data record the trip information of each card holder, such as boarding and alighting stations and the corresponding boarding and alighting time. The Taxi GPS device generated the trajectory data of each taxi during the operation time. The location of a taxi on its trajectory was recorded by a time slice of 60 s. Velocity and operational status were also stored in the taxi GPS data.

In this paper, due to the data limits, smartcard data were collected from October 13 to 19, 2014, and the taxi GPS data was gathered from June 16 to 22, 2014. Those two weeks did not contain public holidays, the structure of the subway network did not change, the climate was similar throughout, and the total number of taxis was the same as usual. Although the two datasets were collected from different times, the fluctuation of subway passenger volume between June and October was under 5%[1], which means the

[1] https://weibo.com/bjsubway

relationship of competition was stable during the study period and the two datasets were applicative in this research. Until October 2014, the subway system in Beijing had 17 lines and 233 stations, according to the GIS dataset.

The OD of each taxi and subway trip were extracted based on the operation status from the GPS trajectory dataset and the station numbers from smartcard data. Passenger travel information including those of taxi and subway travel were prepared with the trip ID, OD location, and boarding and alighting time. Abnormal data were cleaned, such as incomplete records, missing origins or destinations, and trips with illogical travel time and distance. Finally, 30,042,133 subway trips and 1,443,903 taxi trips were used in this study.

B. MAPPING SUBWAY-COMPETING TAXI TRIPS ONTO SUBWAY NETWORK

SCTTs can be defined as taxi trips with ODs within the catchment areas of subway stations that could have been completed by taking the subway [12], [36]. For each SCTT, there is an actual travel route by taxi as well as a potential substitute route by the subway (Fig. 2). To explore the substitute route of each SCTT by subway, the pick-up and drop-off points must be mapped onto the subway network. We assumed that the subway-competing taxi users (SCTUs) will board the subway at the stations nearest to their initial origins and alight the subway at the stations nearest to their initial destinations. Based on the Euclidean distance between a taxi's OD points and the associated stations, the shortest distances of access and egress to/from the subway stations for each SCTU can be determined. Thus, the SCTTs can be identified by comparing the estimated shortest access/egress distances with preset distance thresholds of the catchment areas of subway stations.
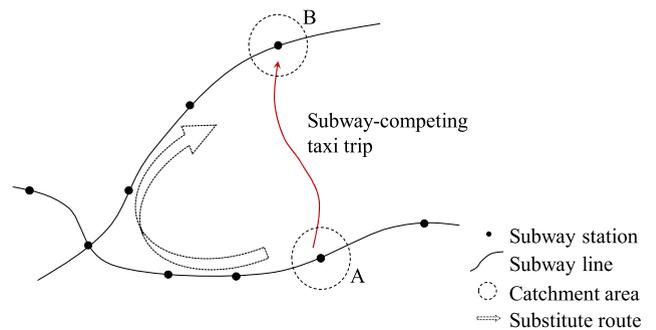

FIGURE 2. Example of subway-competing trip.

We used Python to extract the SCTTs and map them onto the subway network. The identification process is summarized below.

1. For each taxi trip $t(o,d)$ that is represented by the location of the OD, the access and egress distances to each subway station $\mathbf{d}(o,station)$ and $\mathbf{d}(d,station)$ are calculated.
2. The minimum distance $d_{min}(o,station)$ and $d_{min}(d,station)$ are chosen from the dataset; if $d_{min}(o,station)$

$\leq d_{threshold}$ and $d_{\min}(\text{d,station}) \leq d_{threshold}$, then the taxi trip belongs to the subway-competing trip cluster.

3. For each subway-competing trip *sctt(o,d)*, we have the actual pick-up and drop-off location from taxi dataset and mapped boarding and alighting station in subway system. This SCTT can be represented as *sctt(o,d,station_o,station_d)*.

Following the above steps, each pick-up or drop-off point of one taxi trip can be only allocated for a certain pair of stations. By mapping the SCTTs onto the subway network, their alternative subway stations were determined, and the substituted subway route was found based on the OD stations in the subway network.

C. IDENTIFYING BASIC TRAVEL PATTERN BY NMF

To explore the latent travel patterns of the SCTTs, we first constructed a matrix of the OD volume over time based on the mapping result, and then decomposed the demand matrix. We use *d* to denote the number of days, *h* to denote the number of time slots during the day, *n* to denote the number of subway stations, and $m = n \times n$ to represent the total number of station pairs in the subway network. According to the mapping results, we described the directed OD demand of an SCTT by alternative subway station pairs instead of the actual location. Let **V** be the OD demand matrix in the subway network. The element in **V** is represented as $v_{ij,t}$, which is the number of SCTTs on station pair *ij* $(i,j \in [1, n])$ of the substitute subway route in time slot t $(t \in [1, h])$. The demand on the station pair is the sum of passenger volumes in the same time period of each day in (1). The entire OD demand can be calculated by (2).

$$v_{ij,t} = \sum_{d=1}^{d} v_{ij}^{t} \tag{1}$$

$$V_{m \times h} = \begin{pmatrix} v_{11,1} & \cdots & v_{11,h} \\ \vdots & \ddots & \vdots \\ v_{nn,1} & \cdots & v_{nn,h} \end{pmatrix}_{m \times h}$$

$$= \begin{pmatrix} \sum_{d=1}^{d} v_{11}^{1} & \sum_{d=1}^{d} v_{11}^{2} & \cdots & \sum_{d=1}^{d} v_{11}^{h-1} & \sum_{d=1}^{d} v_{11}^{h} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{d=1}^{d} v_{1n}^{1} & \sum_{d=1}^{d} v_{1n}^{2} & \cdots & \sum_{d=1}^{d} v_{1n}^{h-1} & \sum_{d=1}^{d} v_{1n}^{h} \\ \sum_{d=1}^{d} v_{21}^{1} & \sum_{d=1}^{d} v_{21}^{2} & \cdots & \sum_{d=1}^{d} v_{21}^{h-1} & \sum_{d=1}^{d} v_{21}^{h} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_{d=1}^{d} v_{nn}^{1} & \sum_{d=1}^{d} v_{nn}^{2} & \cdots & \sum_{d=1}^{d} v_{nn}^{h-1} & \sum_{d=1}^{d} v_{nn}^{h} \end{pmatrix} \tag{2}$$

Using NMF, the OD demand **V** can be factorized into two lower-ranked non-negative matrices in (3). Each row of **V** can be computed by (4), where $\mathbf{V_i}$ is the *i-th* row vector of matrix **V**, $\mathbf{S_i}$ is the *i-th* row vector of matrix **S**, and $\mathbf{B_i}$ is a *1 × h*

vector.

$$\mathbf{V}_{m \times h} \approx \mathbf{S}_{m \times k} \mathbf{B}_{k \times h} \quad (k \ll \min(m, h)) \tag{3}$$

$$\mathbf{V}_i = \mathbf{S}_i \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_k \end{bmatrix} \tag{4}$$

It is obvious that each row of **V** can be described by the linear combination of $\mathbf{B_i}$ with the coefficients in **S**. Their temporal distributions are revealed in **B**. Moreover, the dominant basic pattern of each station pair can be determined from matrix **S**. Let us take an example to explain the practical meaning of the result of matrix decomposition. If we assume that three basic patterns were identified, then $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3]^{\mathbf{T}}$, and the OD demand of station A to station B was decomposed as follows: $\mathbf{V_{AB}} = \mathbf{S_{AB}B} = [0.1, 0.2, 0.3][\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3]^{\mathbf{T}}$. The coefficient of $\mathbf{B_3}$ in $\mathbf{S_{AB}}$ is the highest; this means the basic pattern-3 is the primary feature vector, and passengers on station pair AB are likely to travel according to pattern-3.

Before decomposing the SCTT demand matrix, the number of basic patterns *k* must be predetermined. Existing studies on the travel pattern of taxi trips determined *k* through experiments with different initial conditions and found that when *k* = 3, the decomposing results can be stable and explicable [17], [29]. Since the SCTT is a subset of taxi trips, we followed the above approach by choosing *k* from 2 to 4.

D. MODELING SUBWAY COMPETITIVENESS

The competition between subways and taxis cannot be fully understood only from the spatiotemporal travel pattern of taxi users. Hence, we should quantify competitiveness. In this study, the definition of competitiveness is the ability of competitors to attract more users toward them in the same market. In this paper, the two competitors are the two traffic modes subway and taxi. The market is the total number of passengers who can travel by subways and taxis. A new index was designed to measure the intensity of competition at the station level based on the actual choices between the two modes, which can be the most intuitive and veracious evaluation index. The mode that is chosen by more passengers will have a bigger share of the market, and consequently greater competition. To attract more taxi passengers to subways, we propose the competitiveness index from the perspective of the subway system. Therefore, the index is calculated as the ratio of subway passengers and the sum of subway passengers and the associated SCTUs with a competitive relationship. The function is as follows in (5):

$$Com = \frac{V(S)}{V(S) + V(SCTT)} \tag{5}$$

where *Com* is the subway competitiveness, *V(S)* is the subway ridership, and *V(SCTT)* is the number of SCTTs.

It is obvious that for the entire passenger demand market, the subway has steeper competition than the taxi because of

the huge capacity of the subway system. However, at the station level, this might be different. Subway competition for each station pair in the network should be explored by the above function. Moreover, the average competitiveness of all station pairs was regarded as the competitiveness of the subway network. The daily and hourly competitiveness of the entire network are expressed in (6) and (7), respectively.

$$Com_{avg}^d = \frac{1}{m} \sum_{ij=1}^{m} \frac{\sum_{t=1}^{h} v_{ij}^t(s)}{\sum_{t=1}^{h} v_{ij}^t(s) + \sum_{t=1}^{h} v_{ij}^t(sctt)} \tag{6}$$

$$Com_{avg}^t = \frac{1}{m} \sum_{ij=1}^{m} \frac{v_{ij}^t(s)}{v_{ij}^t(s) + v_{ij}^t(sctt)} \tag{7}$$

The value of subway competitiveness is between 0 and 1. The greater the value, the steeper the competition by the subway. When $Com = 1$, subway competition is the highest and the number of SCTTs is 0, which means all passengers who could finish trips by either subway or taxi chose subway; when $Com = 0$, the subway ridership is 0 and the subway competition is the lowest.

IV. RESULTS AND ANALYSIS

Based on prior knowledge of the catchment of subways [46], we adopted 1 km as the distance threshold ($d_{threshold}$) when extracting the SCTTs. As a result, there were a total of 679,202 taxi trips belonging to the SCTT group, which accounted for 47.04% of all taxi trips in a studied week; this ratio was highly stable for each day of the week (with differences less than 2%) (Fig. 3). From Wednesday to Sunday, SCTT ridership saw a slight reduction in line with the total taxi trips, and it appeared that fewer people would travel by taxi on weekends. Compared with the other two types of trips, i.e., subway-extending and subway-complementing taxi trips, the SCTT was the most important component. More specifically, almost half the taxi users could have finished their trips by subway. These taxi users were crucial as the major population of target customers to adopt the subway system.
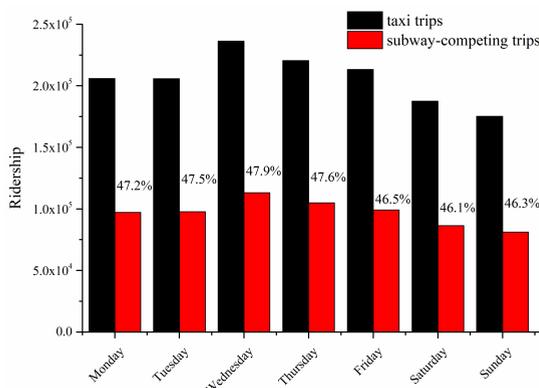
FIGURE 3. Ratio of subway-competing taxi trips to total taxi trips.

A. TRAVEL PATTERN OF THE SUBWAY-COMPETING TAXI TRIP

Analysis of travel patterns of SCTUs can help us better understand the competition and provide useful insights on drawing taxi users to subways. We depicted their travel patterns from the temporal, spatial, and collective aspects, respectively.

1) TEMPORAL PATTERN

Fig. 4 illustrates the temporal distributions of total taxi trips and SCTTs. The variation trend of SCTTs is similar to that of the total taxi trips on both weekdays and weekends. On weekdays, there were three peaks over time for the number of taxis and SCTTs. Subway-competing ridership increased from 4 am, when the number was the lowest, and then reached the first peak between 09:00 and 11:00, which was later than the generally acknowledged morning rush hours in Beijing (07:00–09:00). The second peak appeared between 13:00 and 14:00. In the evening rush hours (17:00–19:00), the demand for the SCTT dropped and then rose to the third peak at 21:00. We inferred that the time periods of SCTT demand peaks were later than the rush hours and varied with the subway demand distribution which with two intensive peaks during peak hours of morning and evening in Beijing [38], [47]. Hence, we supposed that the main travel purpose of subway-competing passengers is not to commute; this point could also be demonstrated by the ratio distribution of SCTT to the total taxi trips in Fig. 5. In the two rush hour periods, the ratio of SCTUs was at a lower level (41–45%) on weekdays, while during off-peak hours, the SCTT demand accounted for a higher proportion of approximately 0.5 and reached peaks during 04:00–05:00, 13:00–14:00 and 23:00–24:00, respectively. Competing passengers accounted for the majority of taxi users during the off-peak time where business, entertainment, and leisure activities were relatively dominated.
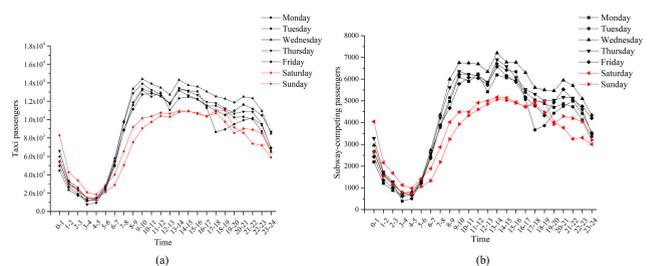
FIGURE 4. Temporal distribution of all taxi trips (a) and SCTTs (b) by hour for one week.

In general, there were fewer SCTTs on weekends than on weekdays, except before 04:00. Whereas the ratio of SCTT was lower on weekends than on weekdays only during daytime (08:00—18:00) on weekends; the value did not change in the morning and the evening. People may have taken fewer SCTTs for business purposes, which was usually more on weekdays.

From the temporal distribution of the SCTT demand and the proportion on weekdays and weekends, we supposed that
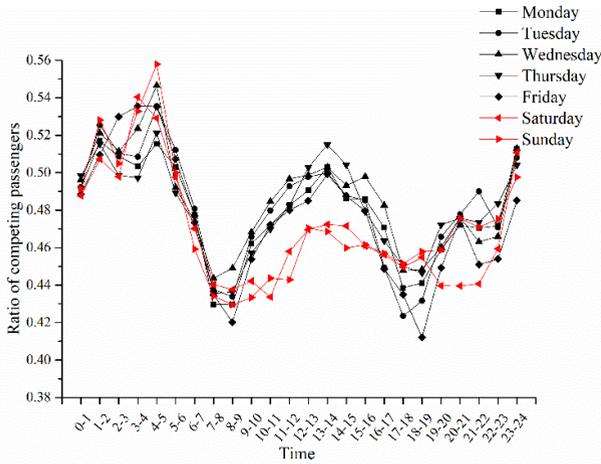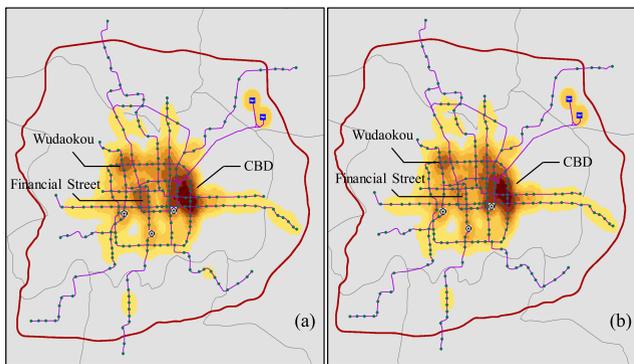
FIGURE 5. Ratio of subway-competing taxi trips to total taxi trips by hour for one week.

the regular commuter is not the main part of SCTUs. The SCTU demand and proportion were greater during off-peak time. Besides, the SCTT percentage was high during closure time (05:00-23:00) of the subway system in Beijing.

2) SPATIAL PATTERNS

For analyzing spatial features, we calculated the density of SCTTs and identified the popular substitute station pairs mapped onto the subway network. Only hotspots of origins on workdays and weekends are shown in Fig. 6(a) and (b), considering the destination distribution was similar to the origin distribution. The darker red color represents the higher density of the origins.


FIGURE 6. Origin spatial distribution of subway-competing taxi trips on workdays.

The spatial distribution of hotspots on workdays and weekends was similar. SCTTs are concentrated in the center of the subway network, which is also the core of Beijing. At the terminals of lines, except the airport express, few SCTTs were generated. The most popular place for subway-competing passengers was the CBD area, which is in the east of the city center. The northwest hotspot is in the Wudaokou area close to many universities and research institutes. It is also the commercial center of the Haidian district. The hotspot in the west of CBD is the Beijing Financial Street. The spatial

distribution of the SCTT revealed that a large collection of citizens would choose taxi rather than subway for business trips, even if the subway station were accessible by walk. Moreover, transportation hubs such as airports and high-speed railway stations were also the frequently visited places for SCTUs; these travelers might pursue the convenience of door-to-door and the comfort of independent spaces brought in by taxis.

To clarify the spatial mobility of SCTUs, we chose OD pairs with a larger passenger flow (top 1%, 287 OD pairs on workdays and 218 OD pairs on weekends), and mapped them onto the subway network as in Fig. 7. In these high-frequency OD pairs, we found that a few SCTUs not only depart (or arrive) at the hotspot areas, but also migrate within the communities of CBD, Wudaokou, and Financial Street, respectively. A new community Olympic park, which was not a hotspot of either pick-up or drop-off points, was revealed. Additionally, there were a few longer-distance SCTTs from (to) transportation hubs including high-speed railway stations and airports. Few trips were observed in other regions.
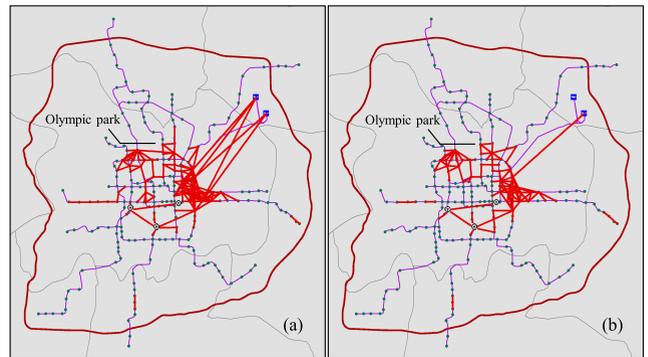
FIGURE 7. High-frequency OD pairs of subway-competing taxi trips.

3) BASIC TRAVEL PATTERN

To better understand the SCTUs, we applied the non-negative matrix factorization method and explored the basic patterns of the spatiotemporal demand, as described in Section 3. The OD points of SCTT were mapped onto 233(n) subway stations, and a total of 233×233 ($m = n \times n$) station pairs were obtained. A day was split into 24 (h) intervals. After preprocessing, a $54,289 \times 24$ ($m \times h$) demand matrix was prepared.

When using this method, we experimented parameter k from 2 to 4. The results are displayed in Fig. 8. When $k = 3$, the decomposition result was more stable and explicable; this is consistent with the findings of previous studies on taxi trips [7], [29], [37]. When $k = 2$, the result explaining the travel patterns was coarse-grained. When $k = 4$, the basic pattern was superfluous to perceive SCTTs. Hence, we took $k = 3$ as the parameter of this method.

The three basic patterns of SCTT are: business or entertainment traveling between workplace and leisure place in the daytime (Pattern1), commuting from home to workplace (Pattern2), and night entertainment from entertainment
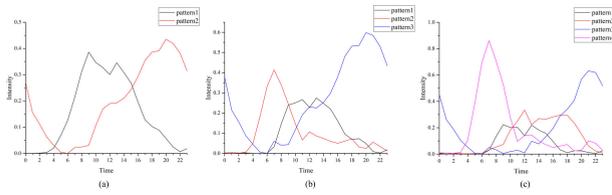
FIGURE 8. The basic patterns for SCTT, where in subfigure (a) k = 2, in subfigure (b) k = 3, and in subfigure (c) k = 4

venues to home (Pattern3). We found that the basic pattern of SCTT is diverse according to the findings of a previous study on total taxi users in Beijing [7], although the SCTT is a subset of taxi trips. Another pattern was found from the total taxi trips, i.e., workplace to home in the evening rush hours, which was not obtained from the SCTT. Moreover, the pattern of trips from the workplace to leisure place in the daytime (Pattern1) was not observed from the total taxi trips.

Based on the above method, we performed a case study that identifies the specific dominating basic pattern for each station pair on workdays. Pattern1 dominated the least number of station pairs, accounting for only 21.24%. 29.02% of station pairs generating trips mainly presented pattern2; and pattern3 prevailed between most station pairs, accounting for approximately 48.95%. According to the different ratios, subway-competing passengers on workdays preferred to head home from the entertainment venues by taxi at night. We then clustered the high-frequency station pairs into the three basic patterns (Fig. 9). Once the station pairs involve transportation hubs (airport and high-speed railway stations), more subway-competing passengers follow pattern2. This meant that subway-competing passengers most likely ply to train stations or airports by taxis in the early morning. Within the CBD and Wudaokou commercial areas, SCTTs were frequently taken at night for entertainment or because of working late hours. According to the specific dominant basic pattern of each station pair, elaborate policies considering the specific time and location can be put forward to draw taxi users to the subway system.
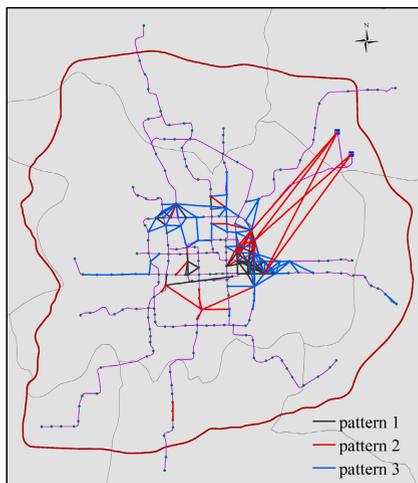

FIGURE 9. Spatial distribution of dominant basic pattern of station pairs with high-frequency SCTTs on workdays.

B. SPATIOTEMPORAL ANALYSIS OF COMPETITIVENESS

For each subway station pair, the ratio of subway passengers to the total number of passengers (subway and taxi users) is regarded as the competitiveness. The actual travel demand was determined for each subway station pair from the smartcard data, since tap-in and tap-out stations are recorded. Meanwhile, pick-up and drop-off points of SCTTs were also mapped onto the nearest stations.

1) TEMPORAL DISTRIBUTION

We calculated the number of subway passengers and SCTUs per hour for each station pair, and then obtained the subway competitiveness matrix $\mathbf{Com}_{54289 \times 24}$. The element $com_{ij,t}$ in matrix $\mathbf{C}$ represents the competitiveness value of station pair ij in time slot t.

To demonstrate the temporal heterogeneity of competition from the perspective of subway network, the distributions of two different time granularities were provided, i.e., daily competitiveness and hourly competitiveness. Fig. 10(a) depicts the temporal distribution of the entire subway network's competitiveness for each day in the studied week. The values were stable in the range of 0.4 to 0.5, which meant the subway services were not attractive for competing passengers. The index peaked on Fridays, since some residents or students from suburban areas worked and studied in the inner city on weekdays and went back home to spend the weekends by subway, causing the subway volume to increase. Fig. 10(b) shows the average competitiveness distribution of all station pairs per hour. The subway attraction on weekends was lower than on workdays in the peak period and higher than on workdays in the off-peak period. Subway competitiveness had two peaks at rush hours in Beijing, 07:00 and 18:00, and on weekends it was steadier than on workday during the daytime. Compared with taxis, subways were more attractive to commuters. However, in the off-peak daytime (10:00–16:00), the index value declined rapidly, and the number of SCTUs was larger during that time (Fig. 4(b)). Moreover, subway competitiveness was lower than 0.5 before 06:00 and after 20:00, which indicates there were more taxi users than subway passengers between the corresponding station pairs. Hence, the subway system should increase its competition urgently during the off-peak time.
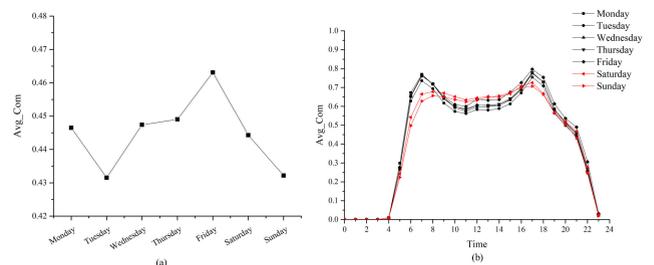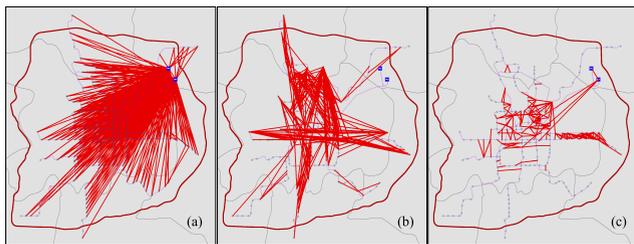

FIGURE 10. Temporal distribution of competitiveness by different time granularity of day (a) and hour (b).

From the temporal characters of subway competitiveness, we conclude that:

1) Competition is dynamic over time in station pair level.
2) Subway has low competition during off-peak time.

2) SPATIAL DISTRIBUTION

In this section, we detail subway competitiveness on workdays in the studied week. Subway competitiveness was 0 when there were no subway passengers on that station pair, and 1 when no SCTUs. Subway competition is worst when the $Com = 0$, all the passengers travel by taxi instead of subways on that station pairs which is shown in Fig. 11(a). Almost all station pairs of zero competitiveness involved the airport. Obviously, passengers preferred to travel to/from airports by taxis rather than the subway. The average travel distance on these station pairs was 33.19 km, which is twice the average distance of subway passengers (15.5 km) and four times the SCTTs (8.5 km). If passengers adopted substitute routes on the subway network, they would have to travel at least once to the airport express. Generally, passengers who went to the airport usually carried more luggage with them. In this case, due to the long distance, multiple transits, and heavy loads, the subway competitiveness was the least on these station pairs, and it was hard to encourage those passengers to take the subway in practice.


FIGURE 11. Spatial distribution of station pairs with competitiveness (a) *Com* = 0, (b) *Com* > 0.5(the top 1% quantile), (c) *Com* ≤ 0.5 on workday.

For most of the station pairs, subway competitiveness index was over 0.5. We only demonstrated the station pairs with 1% highest competitiveness on workdays in Fig. 11(b). In the higher competition group, the average taxi travel distance reached 20.55 km, which meant that the subway may be more attractive to the passengers traveling long distances. These station pairs had a typical feature of converging from the edge of the subway network to the core city. Due to the high cost of housing and rent in Beijing, many residents choose to live in the outer city and commute long distances every day. For these passengers, taxis are too expensive as a daily transportation mode, while the subway may be a better choice.

Except for station pairs with zero and high competitiveness, the rest had less competitiveness (0 < Com ≤ 0.5); these station pairs were the key to attracting more SCTUs to adopt the subway (Fig. 11(c)). The average competitiveness on these station pairs was 0.311; subway passengers made up less than half of the total number of passengers. Most taxi users in this group traveled within the inner city, especially within the commercial areas. However, there were taxi users

who traveled to the airport via the low-competition station pairs. The average distance of taxi trips in this group was 4.31 km, which was about half of the average travel distance of SCTT and one fifth of higher competition group(20.55km).

Furthermore, three typical spatial structure patterns were identified in the group of station pairs with low competitiveness (Fig. 12(a)), i.e., collinearity corridors (CCs), radial communities (RCs), and paralleled subway lines links (PSLLs). CCs are the low-competition corridors shown in Fig. 12(b), which are made up of adjacent station pairs with similar directions. In these corridors, more passengers chose taxis over subways. RCs depicted the radial structures that connect from/to some fixed stations shown in Fig. 12(c). A few low-competition areas were formed around the radial stations. Moreover, the subway network of Beijing has a few paralleled lines close to each other as in Fig. 12(d). Traveling on PSLLs needed to transfer at least twice which was not conventional especially for short trips. Consequently, between these stations, subway had lower competition than taxi.
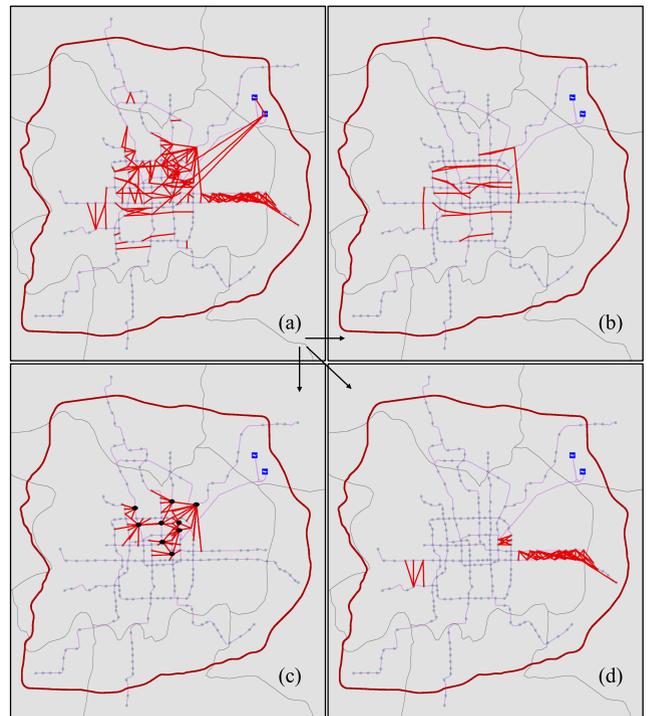
FIGURE 12. Three typical spatial patterns of OD pairs with low competition (a) OD pairs with low competition (b) CCs, (c) RCs, (d) PSLLs.

Characters of SCTTs in high competitiveness group (HCG) were different with those in low competitiveness group (LCG). We summarized basic travel characters of SCTTs by taking taxi and substitute subway in TABLE 1.

In HCG, passengers preferred the subway system as the subway traffic conditions were better than taxi and vice versa. The travel time of subway was shorter than taxi and the average transfer times was just 0.48 which meant more passengers could travel to destinations without transfers by subway. While in LCG, the travel time of subway was longer than taxi and the average transfer times was 1.7 which was about three

TABLE 1. Basic travel characters in LCG and HCG.

Characters	LCG		HCG	
	Taxi	Subway	Taxi	Subway
Avg_travel time(min)	11.53	36.85	41.92	38.72
Avg_congestion time(min)	6.11	0	15.82	0
Avg_min transfer times	0	1.7	0	0.48

times than that in HCG. Moreover, the road traffic congestion was also an external variable to effected subway competition. The result indicated that improving traffic conditions such as shorten travel time and reduce transfer times could benefit for subway competitiveness.

From the spatial distribution of competitionveness, we conclude that:

1) Competition has spatial heterogeneity feature.
2) Subway is a less attractive option for traveling to the airport in Beijing.
3) In the core city and commercial areas, subway competition is low.
4) Three typical configurations, collinearity corridors, radial communities, and paralleled subway lines links are found from low-competition station pairs.

V. DISCUSSION AND CONCLUSION

A. DISCUSSION

In this section we first discuss the advantage of our competitiveness index and then give some suggestions on transportation planning based on the result of travel pattern and competitiveness distribution.

We quantified subway competition to taxis based on actual mode choices results collecting from smart cards and GPS equipment. This method provided more accurate index compared with travel time which was the most used indicator to reflect competitiveness. To prove the better performance of our index, we calculated average travel time difference($\Delta T = T_{subway} - T_{taxi}$) between taxi and subway for each station pair, and drawn the correlation with competitiveness (Com) in Fig.13.

If taking travel time as the competition index, the dots in Fig.13 should only appear within zone B and zone D. However, actually, lots of station pairs located in zone C where the subway travel time was longer but the competitiveness was still in high degree. The phenomenon happened because passengers balanced travel cost not only by travel time but also by some other factors such as comfort, punctuality and even time perception. The mode choice was the comprehensive result of considering all the factors but travel time was only a one-sided indicator. Therefore, the method of quantifying competitiveness in this study was more precise.

Excess competition from taxis to subways leads to road congestion and air pollution, which should be addressed. By analyzing the SCTTs and calculating subway
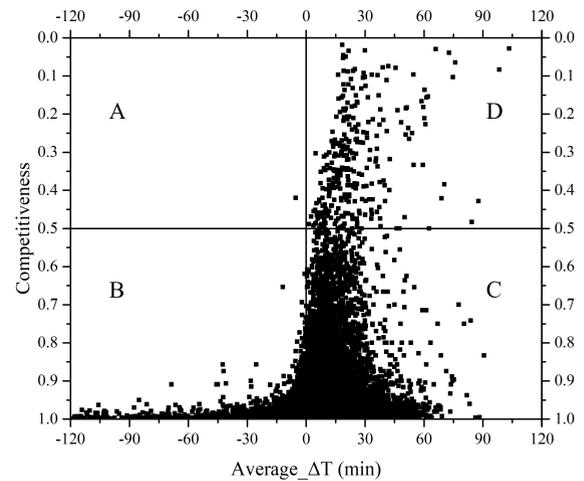
FIGURE 13. Comparison of travel time index and competitiveness index.

competitiveness, we determined the specific constraint condition in terms of each station pair and improved it to enhance the subway competition. We can also determine which station pairs should be promoted on priority. Usually, two strategies are used to tackle the problem of competition, decreasing external driving force, and strengthening internal driving force [21]. In this study, diminishing external driving force reduced the appeal of taxis, such as by increasing the taxi fares based on the associated external cost of road congestion, energy consumption, and environmental pollution. According to the previous study, the internal driving force also has a strong influence on passengers' decision, such as by improving the subway service level for the present and optimizing the planning of subway stations and lines for the future [20]. Next, we discuss how to tackle subway competitiveness according to various competition situations of each station pair.

1) PRIMARY TARGETS FOR THE PRESENT BY INCREASING TAXI FARE

Combining the three basic patterns of the high-frequency station pairs (Fig. 9) and their associated competitiveness value, we identified the subway station pairs with low competitiveness and high demand as the primary targets. Meanwhile, station pairs with low competitiveness and low demand were regarded as the secondary targets. In Fig. 14, 13 station pairs are the primary targets at present on workdays, and therefore should be promoted first. Pattern3, which is indicated by blue lines in the figure, is still the majority among these 13 station pairs, which indicates that passengers at night are essential targets. A dynamic fare scheme over time and location is low-cost and effective for the short term. Fares for station pairs with pattern1 can be increased during daytime, while fares for station pairs with pattern3 should be raised at night.

2) SECONDARY TARGETS FOR THE FUTURE BY INCREASING SUBWAY SERVICE LEVEL

The results of the spatiotemporal analysis of competitiveness have different policy implications that can be used to devise
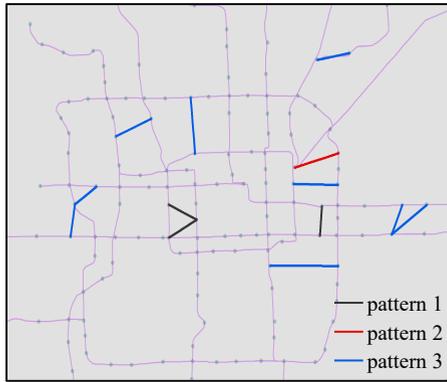
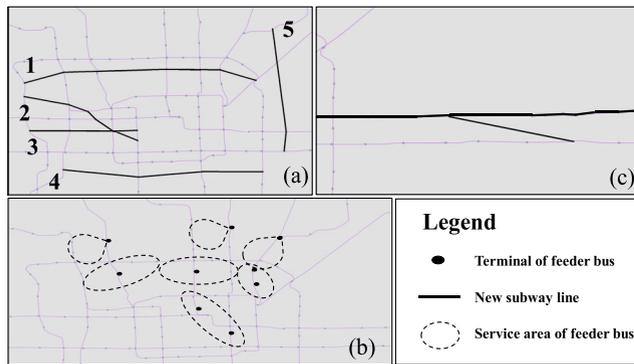**FIGURE 14.** Station pairs with low competition and high SCTT demand.

FIGURE 15. Implementations to increase subway service level: (a) new subway lines, (b) feeder bus, (c) Y-type connection line.

a strategy to encourage taxi users to take the subway. The criterion for devising the strategy is as follows. New subway lines should be designed based on the result of low competitiveness CC in Fig. 12(b). By considering the subway features of high transportation capacity and cost, only those CCs with a large and stable SCTT flow should be considered as the location of potential new lines. Besides, new transfer stations should be constructed between new subway lines with exit lines, and the length of each new line should be considered. According to this criterion, we selected five corridors as in Fig. 15(a), which summarizes the requirements to plan the new subway lines. The result was obtained based on data from June 2014 and demonstrated by the fact that lines 7 and 14 ran along the stations in corridor-4 and corridor-5, respectively, in the end of 2014. Moreover, line 12, which was designed on stations in corridor-1, is scheduled to operate by the end of 2021. This series of empirical verification supports the recommendations of this study on subway planning. Therefore, for subway planners, corridor-2 and corridor-3 shall be given top priority when designing following new subway lines.

New subway lines may not be suitable for other low-competitiveness station pairs, such as OD connections with RCs (Fig. 12(c)). Each community has one core station that is frequented by passengers. Because of the above attributes, the feeder bus in Fig. 15(b) is one of the correct ways of conveying the SCTUs from this type of station pair. One terminal of the feeder bus should be the core station, and the

other one can be in a transfer station, if possible, which will reduce the transfer time of passengers in the meantime.

Besides making these recommendations, we also suggest adopting more connecting lines between two parallel lines in the layout of a subway network, such as the Y-type, as in Fig. 15(c). Passengers traveling between stations in parallel lines will have to transit twice, which will starkly decrease competition from taxis to the subway especially for short distance trips. Adding a few connecting links can mitigate the transfer situation and will appeal to more passengers. Moreover, between close parallel lines, emerging transit modes like bike-sharing may reduce competition from taxis.

B. CONCLUSIONS

Subway as the dominant public transit to relieve the pressure of road congestion and air pollution is widely acknowledged. Given the high construction cost, governments often launch new subway lines when the population density reaches a certain level. Different from the subway, taxis as another type of public transport system is flexible without limiting passenger volume and location. It can substitute the subway service in the area where no subway is provided. It can also compete with the subway in the core city. The result of excessive competition between taxis and subways is undesirable for the economy. Other emerging online taxi-like services like ride-hailing and ridesharing can also add to the competition between taxis and subways. Therefore, it is necessary to assist authorities in encouraging more taxi users to adopt the subway system in the inner-city area, to mitigate congestion and pollution from urban road traffic. The trip information of taxis and subways passengers can be tracked by GPS and smartcard data, thus enabling to study the travel pattern of SCTUs from individual points of view and the actual degree of competition with the subway at the OD level. Based on multi-source datasets, including taxi GPS data, smartcard data, and GIS data, this study revealed the spatiotemporal travel patterns of subway-competing users and competitiveness between subways and taxis. Particularly, a new index to quantify the degree of competition was proposed. Compared with the existing indices, this index is more sensitive and accurate. Based on the travel pattern of SCTTs and the competitiveness index, we can see that the SCTT has its own characteristics. Many of them are quite different from taxis, such as dominating travel purposes and basic travel patterns. The important findings of this study can be summarized as follows.

1) TRAVEL PATTERN OF SCTT

(1) Competition of taxis with subways cannot be ignored since the effects show on a large scale. SCTTs account for the largest proportion of total taxi trips, approximately up to 50%.

(2) Commuters are not the main part in SCTUs. The proportion of SCTT is at the bottom during the morning and evening rush hours and reaches its peak at noon and night.

(3) SCTUs visit business and entertainment centers frequently. Three hotspots of pick-up and drop-off points were detected, namely the CBD area, Wudaokou, and the Financial Street. Most of the popular OD pairs are within these three areas and present the community character. Besides business areas, trips from/to transportation hubs including high-speed railway stations and airports are also in great demand.

(4) Three basic travel patterns—trips between workplace and leisure place in the daytime (Pattern1), commuting from home to workplace (Pattern2), and night entertainment from entertainment venues to home (Pattern3)—were determined based on the non-negative matrix factorization method. Most station pairs visited by SCTUs following Pattern3.

2) COMPETITIVENESS

(1) The average competitiveness in the studied week is not in the high range. On Fridays, the competitiveness index peaks as the subway volume increases. Subway competition is lower in the peak time and greater in the off-peak time.

(2) All station pairs involving airports have low competition value. Passengers generally do not prefer taking the subway to the airport.

(3) Most station pairs in low-competition groups are in the core city and commercial areas. High-competition station pairs are the combination of the outer station and the central station.

(4) Connections between low-competition station pairs exhibit three typical configurations, collinearity corridors, radial communities, and paralleled subway lines links.

In summary, the latent travel pattern of SCTUs is understood and the spatiotemporal heterogeneity of competitiveness in subway network is determined. According to the findings, suggestions of short-term and long-term planning to promote the subway competition in practice are recommended.

This study can be regarded as the starting point of improving subway competition, and therefore has several limitations. First, the datasets of taxi and subway were not collected in the same period. However, the subway network did not change during that period and the demand of subway passengers remains a stable level, which can guarantee the reliability of the results. Another limitation is the lack of sufficient validation for the proposed policies or suggestions on improving the subway competition. How to implement these policies and allow them to fulfill their roles should be better investigated.

In future studies, more detailed studies based on subway-competing passengers and subway competitiveness should be conducted. First, considering both passenger attributes and traffic conditions based on survey data and trip data, the relevant factors and their exact correlation with the subway competitiveness performances should be determined.

Then, the multimodal transit of taxis and subways should be developed to enhance the accessibility of subway stations and lead to a more sustainable urban transport.

REFERENCES

[1] D. Austin and P. C. Zegras, "Taxicabs as public transportation in Boston, Massachusetts," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2277, no. 1, pp. 65–74, Jan. 2012.

[2] J. M. Casello, "Transit competitiveness in polycentric metropolitan regions," *Transp. Res. A, Policy Pract.*, vol. 41, no. 1, pp. 19–40, Jan. 2007.

[3] F. Carlsson, "Design of stated preference surveys: Is there more to learn from behavioral economics?" *Environ. Resource Econ.*, vol. 46, no. 2, pp. 167–177, Jun. 2010.

[4] R. Cervero, *Paratransit in America: Redefining Mass Transportation*. Westport, CT, USA: Greenwood Publishing Group, 1997.

[5] F. Chen, Z. Yin, Y. Ye, and D. Sun, "Taxi hailing choice behavior and economic benefit analysis of emission reduction based on multi-mode travel big data," *Transp. Policy*, vol. 97, pp. 73–84, Oct. 2020.

[6] X. Cheng, K. Huang, L. Qu, and L. Li, "A cooperative data mining approach for potential urban rail transit demand using probe vehicle trajectories," *IEEE Access*, vol. 8, pp. 24847–24861, 2020.

[7] Y. Dong, S. Wang, L. Li, and Z. Zhang, "An empirical study on travel patterns of Internet based ride-sharing," *Transp. Res. C, Emerg. Technol.*, vol. 86, pp. 1–22, Jan. 2018.

[8] R. B. Ellison and S. Greaves, "Travel time competitiveness of cycling in Sydney, Australia," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2247, no. 1, pp. 99–108, Jan. 2011.

[9] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva, "Visual exploration of big spatio-temporal urban data: A study of New York City Taxi trips," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2149–2158, Dec. 2013.

[10] N. Gendron-Carrier *et al.*, "Subways and urban air pollution," Nat. Bur. Econ. Res., Cambridge, MA, USA, Tech. Rep. w24183, 2018.

[11] H. H. Hochmair, "Spatiotemporal pattern analysis of taxi trips in New York city," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2542, no. 1, pp. 45–56, Jan. 2016.

[12] S. Jiang, W. Guan, Z. He, and L. Yang, "Exploring the intermodal relationship between taxi and subway in Beijing, China," *J. Adv. Transp.*, vol. 2018, pp. 1–14, Jun. 2018.

[13] L. Kattan, A. de Barros, and S. C. Wirasinghe, "Analysis of work trips made by taxi in canadian cities," *J. Adv. Transp.*, vol. 44, no. 1, pp. 11–18, Jan. 2010.

[14] K. Kim, "Exploring the difference between ridership patterns of subway and taxi: Case study in seoul," *J. Transp. Geogr.*, vol. 66, pp. 213–223, Jan. 2018.

[15] S. C. Kwan and J. H. Hashim, "A review on co-benefits of mass public transportation in climate change mitigation," *Sustain. Cities Soc.*, vol. 22, pp. 11–18, Apr. 2016.

[16] A. Lakhina, K. Papagiannaki, M. Crovella, C. Diot, E. D. Kolaczyk, and N. Taft, "Structural analysis of network traffic flows," in *Proc. Joint Int. Conf. Meas. Modeling Comput. Syst.*, 2004, pp. 61–72.

[17] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[18] H. Lee, H.-C. Park, S.-Y. Kho, and D.-K. Kim, "Assessing transit competitiveness in seoul considering actual transit travel times based on smart card data," *J. Transp. Geogr.*, vol. 80, Oct. 2019, Art. no. 102546.

[19] J. Lee and S. Park, "Resource capacitated collective travel planning in spatial databases," *IEEE Access*, vol. 8, pp. 135443–135457, 2020.

[20] L. Li, M. Cao, Y. Bai, and Z. Song, "Analysis of public transportation competitiveness based on potential passenger travel intentions: Case study in shanghai, China," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 4, pp. 823–832, Apr. 2019.

[21] L. Li, Y. Bai, Z. Song, A. Chen, and B. Wu, "Public transportation competitiveness analysis based on current passenger loyalty," *Transp. Res. A, Policy Pract.*, vol. 113, pp. 213–226, Jul. 2018.

[22] M. Li, L. Dong, Z. Shen, W. Lang, and X. Ye, "Examining the interaction of taxi and subway ridership for sustainable urbanization," *Sustainability*, vol. 9, no. 2, p. 242, Feb. 2017.

[23] S. Li, Y. Liu, A.-O. Purevjav, and L. Yang, "Does subway expansion improve air quality?" *J. Environ. Econ. Manage.*, vol. 96, pp. 213–235, Jul. 2019.

[24] H. Liu, X. Chen, Y. Wang, and S. Han, "Vehicle emission and near-road air quality modeling for shanghai, China: Based on global positioning system data from taxis and revised MOVES emission inventory," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2340, no. 1, pp. 38–48, Jan. 2013.

[25] X. Li, J. Kurths, C. Gao, J. Zhang, Z. Wang, and Z. Zhang, "A hybrid algorithm for estimating origin-destination flows," *IEEE Access*, vol. 6, pp. 677–687, 2018.

[26] X. Liu, H. Titheridge, X. Yan, R. Wang, W. Tan, D. Chen, and J. Zhang, "A passenger-to-driver matching model for commuter carpooling: Case study and sensitivity analysis," *Transp. Res. C, Emerg. Technol.*, vol. 117, Aug. 2020, Art. no. 102702.

[27] X. Liu, X. Yan, F. Liu, R. Wang, and Y. Leng, "A trip-specific model for fuel saving estimation and subsidy policy making of carpooling based on empirical data," *Appl. Energy*, vol. 240, pp. 295–311, Apr. 2019.

[28] J. C. Martín, C. Román, J. C. García-Palomares, and J. Gutiérrez, "Spatial analysis of the competitiveness of the high-speed train and air transport: The role of access to terminals in the Madrid–Barcelona corridor," *Transp. Res. A, Policy Pract.*, vol. 69, pp. 392–408, Nov. 2014.

[29] C. Peng, X. Jin, K.-C. Wong, M. Shi, and P. Liò, "Collective human mobility pattern from taxi trips in urban area," *PLoS ONE*, vol. 7, no. 4, Apr. 2012, Art. no. e34487.

[30] R. M. Pendyala, "Challenges and opportunities in advancing activity-based approaches for travel demand analysis," in *Proc. Expanding Sphere Travel Behav. Res., Sel. Papers From 11th Conf. Int. Assoc. Travel Behav. Res., Sel. Papers From 11th Conf. Int. Assoc. Travel Behav. Res.*, 2009, p. 303.

[31] L. Rayle, D. Dai, N. Chan, R. Cervero, and S. Shaheen, "Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco," *Transp. Policy*, vol. 45, pp. 168–178, Jan. 2016.

[32] D. Sun and X. Ding, "Spatiotemporal evolution of ridesourcing markets under the new restriction policy: A case study in Shanghai," *Transp. Res. A, Policy Pract.*, vol. 130, pp. 227–239, Dec. 2019.

[33] D. Sun, K. Zhang, and S. Shen, "Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data," *Transp. Res. D, Transp. Environ.*, vol. 62, pp. 699–714, Jul. 2018.

[34] X. Sun, S. Wandelt, and E. Stumpf, "Competitiveness of on-demand air taxis regarding door-to-door travel time: A race through Europe," *Transp. Res. E, Logistics Transp. Rev.*, vol. 119, pp. 1–18, Nov. 2018.

[35] E. Tonnelier, N. Baskiotis, V. Guigue, and P. Gallinari, "Smart card in public transportation: Designing a analysis system at the human scale," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 1336–1341.

[36] F. Wang and C. L. Ross, "New potential for multimodal connection: Exploring the relationship between taxi and transit in New York City (NYC)," *Transportation*, vol. 46, no. 3, pp. 1051–1072, Jun. 2019.

[37] J. Wang, F. Gao, P. Cui, C. Li, and Z. Xiong, "Discovering urban spatio-temporal structure from time-evolving traffic networks," in *Proc. Asia–Pacific Web Conf.* Cham, Switzerland: Springer, 2014, pp. 93–104.

[38] Y. Wang, Y. Wang, D. Ettema, Z. Mao, S. G. Charlton, and H. Zhou, "Commuter value perceptions in peak avoidance behavior: An empirical study in the Beijing subway system," *Transp. Res. A, Policy Pract.*, vol. 139, pp. 70–84, Sep. 2020.

[39] F. Witlox, "Evaluating bicycle-car transport mode competitiveness in an urban environment. An activity-based approach Frank Witlox & Hans Tindemans," *Editorial Board*, vol. 10, no. 4, pp. 32–42, 2004.

[40] M. Wohl, "The taxi's role in urban America: Today and tomorrow," *Transportation*, vol. 4, no. 2, pp. 143–158, Jun. 1975.

[41] C. Xiong and L. Zhang, "Dynamic travel mode searching and switching analysis considering hidden model preference and behavioral decision processes," *Transportation*, vol. 44, no. 3, pp. 511–532, May 2017.

[42] C. Xiong, X. Chen, X. He, W. Guo, and L. Zhang, "The analysis of dynamic travel mode choice: A heterogeneous hidden Markov approach," *Transportation*, vol. 42, no. 6, pp. 985–1002, Nov. 2015.

[43] Z. Xu *et al.*, "Quantifying the competitiveness of transit relative to taxi with multifaceted data," *Transportmetrica A, Transp. Sci.*, pp. 1–22, 2020, doi: 10.1080/23249935.2020.1770367.

[44] Y. Ye, J. Sun, and J. Luo, "Analyzing spatio-temporal distribution pattern and correlation for Taxi and Metro ridership in Shanghai," *J. Shanghai Jiaotong Univ.*, vol. 24, no. 2, pp. 137–147, Apr. 2019.

[45] N. Yong, S. Ni, S. Shen, P. Chen, and X. Ji, "Uncovering stable and occasional human mobility patterns: A case study of the Beijing subway," *Phys. A, Stat. Mech. Appl.*, vol. 492, pp. 28–38, Feb. 2018.

[46] M. Yue, C. Kang, C. Andris, K. Qin, Y. Liu, and Q. Meng, "Understanding the interplay between bus, metro, and cab ridership dynamics in Shenzhen, China," *Trans. GIS*, vol. 22, no. 3, pp. 855–871, Jun. 2018.

[47] Z. Zhang, H. Fujii, and S. Managi, "How does commuting behavior change due to incentives? An empirical study of the Beijing Subway system," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 24, pp. 17–26, May 2014.

[48] F. Zou, B. Wu, J. Xiong, and L. Li, "Analyzing public transportation competitiveness based on the theory of planned behavior," in *Proc. Saf., Speediness, Intell., Low-Carbon, Innov. (ICTE)*, Oct. 2013, pp. 3139–3147.

RUI WANG is currently pursuing the Ph.D. degree with Beijing Jiaotong University, China. She was a Research Assistance with University College London from 2018 to 2019. Her research interests include public transport planning and policy and traffic big data analysis, especially in the aspects of urban rail transit and taxi data analysis.

FENG CHEN received the B.S., M.A., and Ph.D. degrees in railway engineering from Beijing Jiaotong University, Beijing, China, in 1983, 1990, and 2007, respectively.

He is currently working as a Professor with Beijing Jiaotong University and the China University of Petroleum, Beijing. His research interests include passenger flow management, traffic data mining, and application for urban rail transit. He won the first prize of national science and technology progress as the second accomplisher in 2017.

XIAOBING LIU is currently pursuing the Ph.D. degree in traffic planning and management with Beijing Jiaotong University, China. He is also one member of the MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport in China. He focuses on transportation energy, transportation Policy & Planning, and traffic big data analysis. He has conducted some data-based research in sharing mobility, especially the management and optimization of on-demand ridesharing services.

TAKU FUJIYAMA started the Ph.D. degree from the Center for Transport Studies, UCL, in 2002. He continued his Ph.D. research while he started working as a Research Assistant in 2004. As a Research Assistant, he was involved in various pedestrian-related projects, including designing platform humps for London Underground and investigation into train dwell time for Thameslink. His research interests include design of transport environments and behaviour of users, railway infrastructure, and operation and resilience of national infrastructure and user responses.

• • •