Research on trip chain characteristics of round-trip car-sharing users in China: A case study in Hangzhou City

Ying Hui, Ph.D. Corresponding Author
Assistant Professor,
Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
School of Transportation Engineering,
Tongji University
4800 Cao’an Road, Shanghai, 201804, China
Tel: +86-13636458669
E-mail: huiying@tongji.edu.cn

Mengtao Ding, Master, Candidate
Graduate Research Assistant,
Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
School of Transportation Engineering,
Tongji University
4800 Cao’an Road, Shanghai, 201804, China
Tel: +86-18701752860
E-mail: dmt1992@126.com

Kun Zheng, Senior Student
Graduate Research Assistant
Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
School of Transportation Engineering,
Tongji University
4800 Cao’an Road, Shanghai, 201804, China
Tel: +86-18817599398
E-mail: 18817599398@163.com

Yi’an Liu, Ph.D.
Research Assistant
Chelizi Intelligent Technology Company
Tel: +86-13634198360
E-mail: 544032672@qq.com

Word count: 4,739 words text + 10 tables/figures x 250 words (each) = 7,239 words

Submission Date: August 1st, 2016
ABSTRACT
Car sharing as a means to provide individuals with a short-term access to automobiles to complete a personal trip has been well received in Europe and North America but is still in its infancy in China. Currently, there are few case studies based on operation data to support the importance of car sharing system in urban comprehensive transportation system. In this study, vehicle GPS data of a round-trip car-sharing system in Hangzhou was used to derive trip information and classify the trip chain pattern. The rental behaviors, of which time is within 24 h, are chosen as study samples. Through data preprocessing, 26,085 valid behavior samples are obtained. Several trip chain characteristic indices are selected to describe the characteristics. The Clustering LARge Applications algorithm was used to analyze the clustering of such sized samples. Next, 26,085 samples were classified into five clusters as nonactivity point, multiactivity point, high stop-rate, long-distance, and common patterns; each cluster has significant differences in one or two trip chain characteristics indices. The cluster result reflected that different use patterns exist, and at least in three typical use patterns, shared vehicles are used as private cars to satisfy continuous daily trip demands or even commute demands. The car-sharing system in Hangzhou still presents a multipattern condition. Therefore, the proposed method could facilitate companies in formulating a flexible price strategy and determining valued customers. Moreover, the traffic administration could more deeply understand the position and function of car-sharing mode in an urban comprehensive transportation system, and plan scientific policies.

Keywords: car sharing; GPS data; CLARA clustering algorithm; trip chain characteristics; use pattern;
INTRODUCTION
With the development of the Chinese economy, the popularization of private cars is increasing rapidly along with the growth of urban mobility demand. This has resulted in associated problems such as traffic congestion and parking dilemmas. In some Chinese cities, the government has formulated a policy to control the growth of private cars. Car sharing, which offers mobility and flexibility of private cars and simultaneously eases pressure on public transportation, can reduce private car ownership to a certain extent (1). Based on the experience in North America, a share car can substitute for six to seven private cars on average and reduce members’ total trip mileage (2). However, car sharing is still a new transportation mode for Chinese people. Therefore, reasonable positions and development policies for car sharing must be formulated. Analyses based on operation data can directly describe the usage and deduce the trip demand. This research focuses on the actual trip chain characteristics based on GPS data to provide reference for subsequent researches.

Introduction of the Chefenxiang car sharing system
“Chefenxiang” is the first operational EV car sharing system in China operable since 2010. This is a round-trip car sharing system, implying that the rental vehicle must be returned to the same service station where it was rented. By 2015, the system had more than 1000 users and 80 stations. The rent fare can be charged by hour, mile, or contract server hours. In addition, there is a preferential price policy if the contract hours were from 5:00 p.m. to 9:00 a.m. The users must register and obtain a smart card before using this service. The GPS data facilitated a more comprehensive understanding of a trip’s spatial-temporal characteristics. The following are the main objectives.

- Identifying the typical trip chain characteristics and analyzing their usage patterns.
- Proposing probable trip purposes and usage preference in each use pattern.

The first objective is discussed for the overall understanding of trip chain characteristics of round-trip car-sharing users. The exact knowledge of trip-behavior segmentation explicates the different types of use patterns. It is quite helpful for researchers and government to clearly understand the importance of car sharing in urban transportation system. By analyzing the correlation of the time-dimension feature and use pattern, the time-based accounting strategy or preferential policy can be formulated and the vehicle use efficiency can be enhanced.

LITERATURE REVIEW
Car-sharing system and use pattern
As car sharing have become a widely used transportation mode over the decades with over a million users worldwide, an increasing number of scholars and institutions are focusing on car sharing.

Studies in Europe and North America
Over the years, there have been several case studies and model validations on car-sharing systems in Europe and North America. A study by Kim focused on the car-sharing usage-pattern in New York City, particularly in marginalized neighborhoods. A linear regression models fitted with collected vehicle inventory data and the results revealed that there is high demand on weekdays and weeknights, and car-sharing usage is highly correlated with the total number of vehicles available (3). Morency et al. analyzed the data issued after one year of operation of a car-sharing system in Montreal, and then described and clustered the usage patterns by using a K-means cluster tool (4). Febbraro et al.
analyzed the state of stations to determine if they were oversupplied, undersupplied, or balanced. Further, they grouped stations with similar behavior into zones and charged prices per pair of zones (5).

Modeling studies by investigations or transaction data on car-sharing user’s behaviors are common topic to researches. De modeled the propensity in adhering to a car-sharing system within the random utility framework through starting from a stated preferences survey (6). Cervero used the method of Logit model and Regression models and concluded that 29% of the member gave up using one or more vehicles while they chose bus more frequently in daily trips (7, 8). Habib presented an economic model to jointly forecast membership duration, the decision to become an active member in a particular month, and the frequency of monthly usage of active members based on the membership directory and monthly transaction data of *Communauto Inc* (9).

**Researches in China**

In China, the lack of practical support contributes to the relatively less research on car sharing. Currently, there are two main research aspects of car sharing by Chinese scholars. One aspect focuses on the willingness of Chinese consumers to accept such a system; the Logit model and theory of planned behavior have been used to analyze this, and its impact factors have been discussed (10–12). The other aspect mainly discusses the development prospect and operation strategy of the system (13, 14).

In summary, European and North American countries have abundant research foundation on analyzing characteristics, modeling research, and operation optimization regarding car-sharing systems. Obviously, in many cities, car sharing is an important transportation mode that alleviates the pressure of parking and reducing car ownership. However, in China, car-sharing studies remain at the developmental level and willingness analysis, and lack details of trip characteristics and user’s usage patterns, which are based on operation data.

**GPS data preprocessing and deriving trip information**

Currently, three methods are used to collect trip data in major studies: resident trip survey, revealed preference and stated preference survey and GPS data. Among the three methods, GPS data has higher accuracy and spatial-temporal continuity. Therefore, we used GPS data in this study to collect trip chain characteristics. The method and technology of deriving trip information from GPS data mainly refer to the existing studies and are appropriately modified according to data analysis and experience.

Entrance into poor-signal buildings, tunnels, or districts with exuberant plants or some operation error may cause vehicle GPS data to deviate and show Data Missing and Data Noise (15). Before the research, data preprocessing is necessary. A method based on calculating the average speed of adjacent records has been recommended to inspect the error trajectory point. The speed threshold for distinguishing noise points was set at 200 km/h (16). This method performs well considering large deviation coordinate points but is inefficient in small offset deviations. An improved preprocessing method based on the average peed rate of change is considered. According to the method, the average speed rate of change of adjacent records must satisfy a constraint of mathematical inequality (17).

The previous research on trip identification can be split into two main methods.

One method is based on record gap of flameout. An early research set the record gap threshold to 120, 90, and 60 s; the analysis result indicates that 120s is a reasonable threshold (18). Some other researchers have adopted similar methods with different thresholds (19–21).
The other method is based on still points. By setting a threshold speed that indicates the still point, when the total time added by continuous still points exceeds the threshold, a trip stop point is marked. Stopher et al. combined speed, acceleration, and direction to confirm a still point where the speed is near 0 km/h and the direction is constant, then the time threshold was set to 120 s (22). Du used 1.15 mile/h as the speed threshold (23).

In summary, mining technology has been studied deeply and could provide methodology for this research.

**DATASET**

**Data overview**

The data used in this research is provided by Hangzhou EVnet Co.: the operator of the “Chefenxiang” car sharing system in Hangzhou City. The data contains 31,446 rental records with the corresponding vehicle GPS data from December 2013 to June 2015. The GPS data has abundant spatiotemporal information of shared vehicles. The data return interval is approximately 30 s, even when the vehicle was not rented, the GPS data was also returned to the database. The moment when a user rents a car and opens its door, a unique order ID is automatically generated. The data review confirmed that it is sufficiently accurate for the research. Table 1 lists the GPS data frame, with the implications shown in the Explanation column.

<table>
<thead>
<tr>
<th>Column</th>
<th>Data type</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORDERID</td>
<td>CHAR</td>
<td>Identification of one rent process, unique</td>
</tr>
<tr>
<td>CARNUMBER</td>
<td>VARCHAR</td>
<td>License number of vehicle</td>
</tr>
<tr>
<td>ADD_TIME</td>
<td>DATETIME</td>
<td>Return time of data</td>
</tr>
<tr>
<td>RENTED</td>
<td>BOOLEAN</td>
<td>1=&quot;been rented&quot; 0=&quot;for rented&quot;</td>
</tr>
<tr>
<td>SPEED</td>
<td>INT</td>
<td>Instantaneous speed of return time</td>
</tr>
<tr>
<td>longitude</td>
<td>FLOAT</td>
<td>Longitude of vehicle position</td>
</tr>
<tr>
<td>latitude</td>
<td>FLOAT</td>
<td>Latitude of vehicle position</td>
</tr>
</tbody>
</table>

**Data preprocessing**

As described by previous studies, the GPS data usually shows different deviations. To increase the accuracy of this research, the data was preprocessed. By referring to the GPS preprocessing method in a previous study, we completed preprocessing based on average velocity and average speed rate of change as follows:

- **Step1**

  The first record must be determined to be correct by calculating the distance \( D \) from the service station. The occupied parking space was considered as the dilemma, and its value was set at \( D < 1 \) km. Tab the accepted order set \( \{ S \} \). Further, the orders not in \( \{ S \} \) were excluded and tagged as invalid orders.

- **Step2**

  For all the orders in \( \{ S \} \), the average velocity between adjacent records are calculated as follows.

\[
\Delta T_{(i, \ i+1)} = add\_time_{i+1} - add\_time_i \tag{1}
\]

\[
\bar{v}_{(i, \ i+1)} = \frac{D_{(i, \ i+1)}}{\Delta T_{(i, \ i+1)}} \tag{2}
\]
where

\[ D(i, i+1) \] is the distance between two adjacent records

If \( \bar{v}(i, i+1) > 200 \text{ km/h} \) (Speed threshold), then the \( i+1 \) record is considered as the noise point.

Next, remove the \( i+1 \) record, use the \( i+2 \) record, and recalculate.

- **Step3**

Ensure that \( \forall i, \ \bar{v}(i, i+1) \) is correct, then test whether \( \bar{v}(i, i+1) \) satisfies the inequality (17) as follows:

\[
\frac{|\bar{v}(i+1,i+2)-\bar{v}(i,i+1)|}{(T(i+1,i+2)+T(i,i+1))}\leq 10 \tag{3}
\]

After preprocessing, the total number of valid orders was 28,553.

To emphasize the feature as an urban traffic mode for daily trips and to distinguish it with the conventional car rental mode, we chose the orders with time less than 24 h for the research. In summary, we obtained a sample set of 26,085 orders.

**CONCEPTS AND METHODOLOGY**

**Concepts of trip chain index**

Trip chain usually reflects the continuous trip features within a period (usually an active day is chosen between leaving and returning home), including spatial-temporal feature as trip time, trip distance and traffic feature as traffic mode, and trip purpose. The vehicle GPS data has highly accurate trip trajectory; however, only the information of the trips in which shared vehicles were used is available, and is thus inaccurate for analyzing the trip purpose. In this research, several available trip chain characteristic indices are chosen to describe and analyze users’ trip chain constitution:

- **Number of activity points**: The number of activity points in a rent cycle.
- **Trip chain radius**: The linear distance between a service station and the farthest activity point from the station

\[
R = \max \{r_{stop}\} \quad |i = 1,2, \ldots, n| \tag{4}
\]

- **Trip chain start time**: The start time of the first trip
- **Trip chain end time**: The end time of the last trip

- **Longest stop time**:

\[
T_l = \max \{|\Delta T_{stop}\} \mid i = 1,2, \ldots, n\} \tag{5}
\]

- **Total trip time**: The total trip time in a rent cycle.

\[
T_{trip} = \sum_i \Delta T_{trip} \quad \tag{6}
\]

- **Stop-trip time rate (t)**: The ratio of the longest stop time and the total trip time. This index reflects the time relative proportion of the main activity and trip time.

\[
t = \frac{T_l}{T_{trip}} \quad \tag{7}
\]

**Algorithm of deriving trip information**

The process of algorithm

The algorithm is based on the still point-based method. According to the experience, we set the
stop-time threshold to 120 s and the speed threshold to 2 km/h. Fig. 1 presents the flow diagram of the algorithm.

Fig. 1 The flow diagram of algorithm.

The stop-time threshold of 120 s works well in some American and European cities; however, the traffic environment may be different in Chinese cities such as Hangzhou. The stop-time threshold must be tested using a density plot, which is presented with a semilogarithmic coordinate system, as shown in Fig. 3.
The density plot of stop time (round-off logarithmic abscissa; unit: hour)

Fig. 2 shows that the peak value corresponds to 120 s. A reasonable speculation about this value is that a part of stop points resulted from traffic factors, such as traffic congestion and queuing up, and the other stop points resulted from non-traffic factors in the neighborhood (120 s). To balance this, we reset the stop-time threshold to 180 s, and all the stop points below 180 s will be considered in the trip process.

The density plot of stop time shows an inflection point near the point of 720 s. This may be because the proportion of a stop type mode increases. Thus, we set 720 as the threshold of the activity point.

Then, we have two concepts:

- **Short stay point**
  This type of stop point usually occurs because of meetings, transfers, goods transportations, refueling, or traffic factors such as queuing up and traffic control. In this research, the stop time of a short stay point is 180–720 s.

- **Activity point**
  Usually, stop-time length of the stop points based on work, activities, leisure differ but are generally longer than a short stay point. In this research, the stop time of an activity point is longer than 720 s. Here, “activity” refers to a generalized stop, that is, if the vehicle is stationary for an entire night or a long time, it is also considered as an activity point.

**Combination of short stay point**

The short stay point is likely to be an additional purpose for a certain travel activity. This activity may be an initiative choice, such as meeting and refueling, and is likely to be passive, for example, queuing up for waiting. Further, being stationary in some additional situations can affect the calculated time of trips.

To reduce the effect of some short stay points, such as queuing and meeting, the combination of short-stay point must be studied. According to the research, there are several trips whose main purpose include transfers, meeting, or some short activities. Thus, we cannot simply merge all the short stay points. We introduce a combination rule, that is, all the short stay points that meet the following two characteristics should be included in the travel rather than a travel destination:
Rule 1: The short stay point whose distance from the last or next stop point is less than 1 km.

Rule 2: The trip time between the short stay point and the next or last stop point is less than 500 s.

Fig. 3 shows the schematic of the short stay point combination.

Fig 3. Schematic of merging rule of a short stay point

After combining the eligible short stay points, we acquired the detailed trip feature including the longitude and latitude of each stop point, trip start and end times, and trip distance.

Clustering tool: CLARA Algorithm

To conclude the typical constitution of a trip chain, a clustering method is used. Clustering analysis has a variety of algorithms. Kaufman and Rousseeuw (24) developed an algorithm especially adapted to large data sets: Clustering LARge Applications (CLARA). In this algorithm, partitioning around medoid (PAM) clustering is applied on data subsets of fixed sizes to convert the overall time and storage requirements into a linear rather than a quadratic function of the total number of objects, thus economizing on the computational time. A standard partitioning method directs its main computational effort for searching among a large number of subsets of k objects ($C_n^k$ possible subsets) for a subset yielding a satisfactory, locally optimal clustering. With the increase in the value of n, the number of subsets increases dramatically; for a fixed k, the rate of increase is in the order of the kth power of n. Another factor with the same effect is the storage requirement, which makes the number of memory locations less dependent on the number of objects, of which it is a quadratic function in the PAM algorithm.

Considering the date scale of this research, the CLARA algorithm is applicable. All multivariate analyses were conducted using R language for statistical computation.

RESULTS AND DISCUSSION

Overview of the trip chain model

As the trip information was mined from the GPS data, all trip chain samples are classified into three models according to the characteristics of the activity point. Fig. 5 shows the model
expressions and statistical results.

**Model I** the trip chain that doesn't have stop point identified  
**Model II** the trip chain that only has short-stay point identified  
**Model III** the trip chain that has activity point identified

**Fig. 4 Trip chain model description and statistics**
According to Fig. 4, 91% of car-sharing trip behaviors comprise at least one activity point, and only 8% comprise short stay points. The remaining 1% has no identified stop points. As aforementioned, several short stay points are due to some trip purpose such as meeting, picking up someone, or goods transportation. If these are the main trip purposes, the corresponding trip chain model is likely to be Model I or II. As the same pattern is reflected for users whose main trip purpose does not involve remaining at a place for a long time, Models I and II are merged into one type of trip chain called type I, and another model, that is, Model III, is marked as type II.

**Analysis of type I trip chain: Nonactivity point pattern**
It is easy to infer that the trip time will account for most of the proportion of the total usage time in type I trip chain. Fig. 5 reflects the probability density distribution of trip chain radius, trip time, and total trip distance.

**Fig. 5 Density plots of (a) Trip chain radius, (b) Total trip distance, and (c) Trip time (type I)**
The plot shows that the maximum proportion of trip chain radius is within 15 km, total trip distance is within 30 km, and trip time is within 2 h. After analyzing the distribution characteristics, we can
conclude that in this use pattern reflected by type I, the users present the trip-feature type as mid-short distance trip and short rent time. Considering that the distribution width of the total trip distance is quite close to the double width of the trip radius, we can infer that probably only one round trip occurred in the rent cycle.

Cluster analysis of type II trip chain

The constitution of type I is relatively simple because of its existing no-activity point; however, it accounts for a low proportion, indicating that it is not the main use pattern for car-sharing users. As shown in Fig. 4, 91% of rent behavior consists of activity points (Type II), and the constitution is variable. Different trip chain characteristics reflect different use patterns. The method of analyzing clusters is efficient in classifying samples that do not have a clear demarcation of the group border. By mining clusters of samples with similar characteristics, we can describe and analyze the group characteristics but not each sample. In this study, the CLARA algorithm was used to complete the cluster analysis.

The first step of cluster analysis involves the determination of variables, and correlation of each variable could influence the cluster result. Therefore, it is preferable to use uncorrelated variables. We choose five indices, that is, the number of activity points, longest stop time, total trip distance, total trip time, and trip chain radius to test the correlation level by using the Pearson correlation coefficient. According to the result of correlation analysis, the total trip distance sufficiently correlates to the total trip time and trip chain radius. By clearly distinguishing each pattern, it is more appropriate to use the ratio of the longest stop time and total trip time to describe the model of trip chain.

Therefore, we use the number of activity points, trip chain radius (km), and stop-trip time rate for the clustering analysis.

Silhouette coefficients are the testing values in the CLARA algorithm. They are values combined with dissimilarity and isolation, and a higher value signifies a better clustering result. Reflecting by the Silhouette coefficients, the result indicates that clustering into four groups produces the best result.

The clustering process categorizes 23,657 cases into four clusters. Table 2(a) shows that the largest (cluster 3) and smallest cluster (cluster 4) contain 59.1% and 4.4% of the cases, respectively. Table 2(b) shows the cluster features. As shown, members in clusters 1 and 3 have a higher similarity, indicating that over 80% usage patterns are similar. In contrast, there are outliers in clusters 2 and 4, which match uncommon use patterns. As the isolation reflects, cluster 2 is considerably far away from the other three clusters, implying that the corresponding usage pattern is quite different.

<table>
<thead>
<tr>
<th>Cluster size (% of the 23657 cases)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>5739 (24.3%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2894 (12.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13993 (59.1%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1031 (4.4%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cluster centers</th>
<th>Number of activity points</th>
<th>Trip chain radius (km)</th>
<th>Stop-trip Time rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster centers</td>
<td>5</td>
<td>12</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
<td>10.89</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>61</td>
<td>0.97</td>
</tr>
</tbody>
</table>

TABLE 2 (a) Clustering result

TABLE 2 (b) Clustering feature index
Cluster profiling and initial interpretation

By comparing the distribution of cluster variables, salient characteristics of each cluster could be shown in a plot (Fig. 6(a)). In addition, we illustrated the density plot of the longest stop time and proportion on different days (work day, day off, holiday) for an in-depth analysis.
Instructions: In this study, a usage behavior belonging to a day off and holiday implies that the pick time is during 17:00 on Friday or the day before a holiday to 17:00 on Sunday or the final day of a holiday.

The profiling of clusters to summarize the usage pattern of each cluster is as follows.

- **Cluster 1: Multi-activity point pattern**
  Cluster 1 contains 24.3% of the entire sample. This type of trip chain comprises more activity points, larger trip chain radius, and stop-trip rate between 1.5 and 3. The most obvious difference is that cluster 1 has the most activity points (basically more than four). In this pattern, users tend to plan several activities; the activity range could be for nearby areas or farther regions. As there are no waiting or transfer times between continuous activities, use of shared cars is more convenient than using transits or taxis. Fig. 7 (b) shows that a larger proportion of the sample reveals the longest stop time within 6 h. This indicates that the vehicle is most likely to be used in the daytime. The slope of the density curve near 7–9 h changes to positive, indicating that the probability of this part of the sample increases. Considering that the stop time is always considerably long, vehicles are probably rented overnight and used for the entire evening or for a second day. The proportion of cluster 1 slightly increased during the holidays and day offs because the trip chain constitutes more multi-activities according to the activity patterns for no workdays such as shopping, leisure, and entertainment.

- **Cluster 2: High stop-rate pattern**
  Cluster 2 contains 12.2% of the entire sample. This type of trip chain comprises less activity points, smaller trip chain radius, and large stop-trip time rate. The stop rate significantly differs from other clusters because the longest stop time reaches over 5 times of the total trip time. Therefore, it is inferred that users are perhaps dependent on vehicle usage, convenience, and comfort rather than efficiency. Fig. 7 shows two peaks on the curve that correspond to the two types of behavior that are significantly different: the duration of staying. The first peak indicates a usage pattern in which users employ shared cars as backup vehicles or a transport tool that must not be used for a long time. The second peak indicates a usage pattern in which the shared car is rented for an entire day or night according to the time of commuting trip. The usage pattern of cluster 2 is similar with that of private cars and the use efficiency is not considered. The proportion of cluster 2 significantly increased during workdays. This further supports its correlation with the commuting traffic.

- **Cluster 3: Common usage pattern**
  Cluster 3 contains 59.1% of the entire sample. This type of trip chain is the most popular among car sharing users, and comprises less activity points, smaller trip chain radius, and regular stop-trip time rate. In this pattern, users tend to use sharing cars to complete 1–3 activity trips, usually not far from the service station. The rental time may not always be extremely long. This illustrates that cluster 3 is mainly based on some definite purpose or social trip. In addition, it is not in accordance with the time characteristics of commuting trip from the viewpoint of longest stop time or stop trip rate. Convenience, efficiency, and cost may affect whether user chooses the car-sharing service. Moreover, the density curve of the longest stop time reflects that the stop time is concentrated on an interval within 3 h.

- **Cluster 4: Long-distance pattern**
  Cluster 4 contains 4.4% of the entire sample and accounts for a small proportion but has distinctive features. The number of points is more dispersed, and the trip chain radii are much larger than those of other clusters: mostly more than 20 km. In this pattern, the rental vehicles are likely to be...
used for outings or visiting relatives and friends because most travel distances are beyond the scope of Hangzhou City. The low stop-trip time rate indicates that the vehicle usage efficiency is high. In such a distance range, the cost of employing a taxi is much higher than that of car sharing and may not meet the needs at a continuous travel. The proportion of cluster 4 significantly increases during holidays, supporting the fact that the trip purposes comprise outing or visiting relatives and friends. Its usage is more similar to the characteristics of private cars.

**Distribution characteristics of trip chain start and end time**

According to the analysis in the previous section, we determined the existence of five typical trip chain constitutions with significantly different characteristics. Furthermore, the corresponding trip purposes and choice preferences are inferred. To further support the aforementioned analysis, we focused on the distribution characteristics of trip chain start and end time. Through statistical analysis of trip chain start and end time distribution, we can focus on whether peak hours for pickup or return exist and the different distribution characteristics among each cluster.
Fig. 7 Pickup and return time distribution of each cluster

From Fig. 7, the following conclusions are presented:

- **Types I and II: cluster 3**
  Types I and II of cluster 3 have similar distributions that are quite scattered without obvious peak hours and mainly appear during the daytime. It further infers that the trip purposes are variable in the two usage patterns.

- **Clusters 1 and 2**
  Clusters 1 and 2 have the same pickup and return peak hours, that is, 4–6 p.m. and 8–10 a.m., respectively. This part of samples accords with the commuting trip time and the vehicle rent overnight for a long stop time. The high kurtosis of the distribution of cluster 2 shows that its usage pattern has an obvious time dimension feature performed as follows: “picking at dusk and returning the next morning.” In addition, if there are more activity trips (over four trips) except commuting trips, they are clustered into cluster 1 for the distribution to have similar peak hours.
Moreover, the distribution of cluster 1 comprises starting peak hours between 7:00–9:00 a.m.

• **Cluster 4**

The distribution shape of cluster 4 is considerably different from the other clusters. It has peak hours but not outstanding. The start time is usually concentrated toward the morning, whereas the end time is always concentrated toward the evening. It is consistent to the previous analysis in which the main trip purpose could be outing or visiting relatives and friends. In addition, there is another peak hour at dusk. In this type of user behavior, the car is probably rented for a long-distance trip until the next day.

In summary, the analysis of distribution characteristics of the start and end times of trip chain is consistent to the previous clustering result. The usage pattern comprising some special purpose may show a more concentrated characteristic for pickup and return times.

**SUMMARY**

Table 3 summarizes the main results and findings of the case study. Five trip chain clusters reflecting different usage patterns that express how residents use car sharing or the type of trip pattern that is suitable in China are mined from the GPS data.

**TABLE 3 Summary of characteristics of trip chain types**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Type I Non activity point pattern</th>
<th>Type II–1 Multi-activity point pattern</th>
<th>Type II–2 High stop rate pattern</th>
<th>Type II–3 Common use pattern</th>
<th>Type II–4 Long-distance pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion</td>
<td>9.0%</td>
<td>22.1%</td>
<td>11.1%</td>
<td>53.8%</td>
<td>4.0%</td>
</tr>
<tr>
<td>Number of activity point</td>
<td>Non</td>
<td>Most(almost &gt; 4)</td>
<td>Few(almost &lt; 3)</td>
<td>Few(all &lt; 3)</td>
<td>Disperse(main in 1–7)</td>
</tr>
<tr>
<td>Trip chain radius</td>
<td>Obviously skew to small side</td>
<td>Common and skew to small side</td>
<td>Common and skew to small side</td>
<td>Common and skew to side</td>
<td>Long(almost &gt; 25 km)</td>
</tr>
<tr>
<td>Stop time rate</td>
<td>Almost zero</td>
<td>Common</td>
<td>High</td>
<td>Common</td>
<td>Common</td>
</tr>
<tr>
<td>Longest stop time</td>
<td>Almost zero</td>
<td>Mostly within 3 h and a part of 8–10 h</td>
<td>Two peaks(More between 10–14 h and less between 1–3 h)</td>
<td>Mostly short stop (&lt; 3 h)</td>
<td>Mostly within 3 hours and quite smooth</td>
</tr>
<tr>
<td>Start and end time distribution</td>
<td>Disperse(main in daytime)</td>
<td>Conspicuous peak hours of Start time: 16–18 End time: 8–10</td>
<td>Conspicuous peak hours of Start time: 16–18 End time: 8–10</td>
<td>Disperse(main in daytime)</td>
<td>Inconspicuous peak hours Start usually at forenoon and end usually in evening</td>
</tr>
<tr>
<td>Probable trip purpose</td>
<td>Picking up someone, goods transportation etc.</td>
<td>Various trip purpose in a rent cycle (Maybe including commuting)</td>
<td>Commuting (Working or attending school)</td>
<td>Daily life trip or Business trip</td>
<td>Outing or visiting Usually in off day and holiday</td>
</tr>
<tr>
<td>Substitutability and inference of preference</td>
<td>Suit transit or taxi, depend on preference</td>
<td>Multi-trip is more suitable to vehicle than</td>
<td>Prefer to choose car in trip, like a</td>
<td>Suit transit or taxi, depend on preference</td>
<td>Similar to traditional car rental service</td>
</tr>
</tbody>
</table>
CONCLUSION AND FUTURE DIRECTIONS

This study is based on GPS data and determines the trip chain characteristic of users every time they use the car-sharing system. Five typical trip chains are obtained through statistical cluster analysis. The clusters have constitutions, reflecting different usage patterns for potential trip demands. By analyzing the different trip chain characteristics, we could speculate users’ trip purposes and substitutability. The distribution characteristics of a trip chain’s start and end time can support the speculation of trip purpose and behavioral attitude in different usage patterns. The conclusions of this study are summarized as follows.

- GPS data, containing detailed trip information, can accurately describe the trip chain characteristics of a car-sharing user’s renting behavior. The clustering methodology can split the trip chain samples into groups that share similar characteristics.
- According to the trip chain characteristics, a round-trip car-sharing system set in Hangzhou City currently presents a multipattern condition, indicating that its position in the Chinese urban comprehensive transportation system is not distinct. In particular, unlike the experiences in Europe and North America, in addition to the short-term usage pattern for the main purpose of daily or business trips, a nonignorable portion of commuting trip exists. It can be speculated that most car-sharing usage patterns are certainly irreplaceable to other modes of transport. Moreover, Chinese private-car ownership rate is lower than in European and American countries; however, the traffic stress in many big cities is increasing. Under car-ownership restriction policy in some big cities, the cost of owning a private car is quite high. From this research, car sharing can be considered an important supplement of personal motorization level and remitting car ownership demand.

It must be noted that whether for an enterprise’s individual operation management or traffic administration’s scientific decision about car sharing, the demand patterns of users differ with their usage patterns. An enterprise can analyze a user’s demand in different patterns to manage strategies accordingly. The traffic administration can analyze users’ usage patterns to understand the functional position of car sharing. Finally, this study can be the basis of follow-up studies. Further, this study merely discusses the classification of a trip chain and several features of time dimension, while the trip information contains abundant spatial dimensions. Follow-up studies must focus on different types of spatial distributions of trip chains, land use characteristics surrounding activity points, and the trip-chain type constitution of service stations. In addition, the combination of station-planning issues and Point Of Interest data must be discussed. Moreover, the individual usage pattern of users and use classification could be studied deeply.

ACKNOWLEDGEMENT

This research is supported by the National nature science foundation of China (51408430)

Thanks for data support by EVnet Co. Appreciate for assistance from colleagues: C. Qian (Ph.D.) Q. Sun (Master student), C. Liu (Master student)
REFERENCE

19. Pearson D. Global Positioning System(GPS) and travel surveys: Results from the 1997 Austin household survey. 2001


