



Article Trip Extraction of Shared Electric Bikes Based on Multi-Rule-Constrained Homomorphic Linear Clustering Algorithm

Xiaoqian Cheng¹, Chengming Li^{2,*}, Weibing Du¹, Jianming Shen² and Zhaoxin Dai²

- ¹ School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China; cxq@hpu.edu.cn (X.C.); dwb@hpu.edu.cn (W.D.)
- ² Chinese Academy of Surveying and mapping, Beijing 100830, China; shenjm@casm.ac.cn (J.S.); daizx@lreis.ac.cn (Z.D.)
- * Correspondence: cmli@casm.ac.cn

Received: 1 November 2019; Accepted: 24 November 2019; Published: 26 November 2019



Abstract: Trajectory data include rich interactive information of humans. The correct identification of trips is the key to trajectory data mining and its application. A new method, multi-rule-constrained homomorphic linear clustering (MCHLC), is proposed to extract trips from raw trajectory data. From the perspective of the workflow, the MCHLC algorithm consists of three parts. The first part is to form the original sub-trajectory moving/stopping clusters, which are obtained by sequentially clustering trajectory elements of the same motion status. The second part is to determine and revise the motion status of the original sub-trajectory clusters by the speed, time duration, directional constraint, and contextual constraint to construct the stop/move model. The third part is to extract users' trips by filtering the stop/move model using the following rules: distance rule, average speed rule, shortest path rule, and completeness rule, which are related to daily riding experiences. Verification of the new method is carried out with the shared electric bike trajectory data of one week in Tengzhou city, evaluated by three indexes (precision, recall, and F1-score). The experiment shows that the index values of the new algorithm are higher (above 93%) than those of the baseline methods, indicating that the new algorithm is better. Compared to the baseline velocity sequence linear clustering (VSLC) algorithm, the performance of the new algorithm is improved by approximately 10%, mainly owing to two factors, directional constraint and contextual constraint. The better experimental results indicate that the new algorithm is suitable to extract trips from the sparse trajectories of shared e-bikes and other transportation forms, which can provide technical support for urban hotspot detection and hot route identification.

Keywords: shared e-bikes; trajectory; trip extraction; multi-rule constraints; homomorphic linear clustering

1. Introduction

Various forms of trajectory data are collected owing to the popularity of GPS devices and positioning technology. Trajectory data not only include spatio-temporal information of moving objects but also include abundant interactive human-region, human-society, and even person–person information attributes, which reflect the behavior characteristics of individuals or groups [1,2]. It is helpful to understand and optimize urban decisions by mining the potential information of trajectory data. Trajectory cleaning, aimed at identifying or extracting trips from unordered GPS points, is the key to mining the potential information of trajectory data. The correct identification of trips is helpful for understanding and optimizing urban construction, such as bike lane planning [3], energy conservation

assessment [4], road investment assessment [5], urban functional zones identification [6,7], and urban planning [8–10].

Trips contain rich semantic information, which is useful for understanding human mobility behavior and urban systems. For example, trips can reveal urban resident activities' spatial characteristics and the mutual influence of urban functions' spatial layout and resident activities [11]. The origin–destination (O/D) points of trips are special stop points and imply human activities, which are the basic data for detecting urban hotspots [12–14]. To optimize urban transport management, trips combined with road networks can be used to detect hot routes [15–17]. Thus, distinguishing the O/D points is the key to identify trips.

In the existing literature, the identification methods of O/D points were classified into three categories. The first category can be applied to trajectory data containing completeness attributes or uniform GPS information, such as taxi trajectory or smart card data of buses or subways, whose O/D points were obtained directly through the pick-up/drop-off points [18,19] or the smart card transaction locations [20–22]. However, it is usually difficult to acquire attribute information due to privacy protection. Moreover, GPS signals may be lost because of tall buildings or trees, resulting in missing data. To address such problems, other methods were subsequently proposed to circumvent these shortcomings, such as the time-interval-based method, the clustering-based stop and move of trajectory (CB-SMoT) algorithm, and velocity sequence linear clustering (VSLC). The second category simply regarded trajectory data as another form of spatio-temporal data, whose O/D information was obtained based on different criteria, such as duration, speed, distance, directional change, and so on. For example, Krumm and Horvitz [23] identified O/D points based on speed and time interval, where a point was considered as a step when the speed was lower than 2 km/h or the GPS signal was lost for longer than 5 minutes. Du [24] examined the influences of different combinations of the four indexes of the dwell time, distance, speed and direction change, and different parameter threshold settings on the identification of the end points of private cars' journey. Li and Zheng [25] extracted stops based on the dwell time and distance and further evaluated the similarity of users' stay behaviors. Jia Tao [26] improved this method by using multiple time thresholds instead of a single threshold to extract stops, thus providing data support for the detection of residents' activity patterns. However, the potential semantic information of trips was not considered in these methods. To examine the potential semantic information, more and more scholars have focused on research related to the semantic trajectory, such as semantic trajectory modeling, representation, segmentation, analysis, and mining. The semantic trajectory, first introduced by Spaccapeietra [27], consists of a sequence of stopping/moving objects that can be annotated with important geographical semantics. The semantic trajectory can be expressed by the stop/move model, proposed by Alvares [28]. Thus, the third category mainly involved using the stop/move model to describe the trajectory, which was regarded as spatio-temporal data containing semantic information. Considering the wide application of clustering in data mining, scholars proposed different methods based on the combination of the stop/move model and clustering to detect stops. For example, the clustering-based stop and move of trajectory (CB-SMoT) algorithm [29] used the average speed of the neighboring points as the criterion to construct the stop/move model. However, the algorithm cannot identify stops when few points exceeded the speed threshold. The direction-based stop and move of trajectory (DB-SMoT) algorithm, proposed by Rocha [30], used directional changes to detect stops. However, the DB-SMoT algorithm only worked under certain circumstances in which the direction played an essential role, such as the analysis of fishing vessel trajectories. Xiang [31] proposed the sequence-oriented clustering (SOC) algorithm to extract stops from pedestrian trajectories with noise points and visualized the stops by combining the kernel density and the stop index [32]. The results showed that the SOC algorithm was more suitable for dense sampling trajectory data (the sampling interval was less than 1 minute).

Trips can also be identified by trajectory segmentation, i.e., splitting trajectories into homogeneous segments based on some criteria. Trajectory segmentation is an important task in trajectory data processing, as its correctness will largely affect such subsequent analyses as O/D matrix construction, trip purpose identification, and travel mode detection. The methods of trajectory segmentation can be classified into "supervised" and "unsupervised". The criteria used by "supervised" methods for segmenting trajectories are predetermined. The aforementioned algorithms of O/D identification are "supervised", whether the criteria are monotone or non-monotone. The supervised segmentation methods are user-driven, relying on user-defined rules or thresholds. When the segmentation criteria are unknown, the methods are called "unsupervised". Unsupervised algorithms derive the homogeneity of segments based on some cost function, which can avoid any control from the user, but may lack semantics (W-K means) [33] or are time-consuming (GRASP-UTS) [34]. To combine the benefits of both supervised and unsupervised strategies, Junior [35] first proposed a semi-supervised approach RGRASP-SemTS (Reactive Greedy Randomized Adaptive Search Procedure for Semantic semi-supervised Trajectory Segmentation), which exploited labeled and unlabeled data to achieve homogeneity of segments using a cost function based on the minimum description length (MDL) principle. RGRASP-SemTS can obtain more than 50 features to enrich trajectory data, which is useful for semantic trajectory segmentation. To verify algorithm performance, RGRASP-SemTS was performed with the Atlantic hurricane trajectory dataset and grey seals trajectory dataset.

In summary, few works in the literature of trip identification are related to the sparse trajectory data of vehicles. Therefore, it is worth studying and examining how to correctly extract trips from sparse trajectory data. Shared electric bikes (e-bikes) are selected as an example, as e-bikes can easily travel along wide roads and narrow alleys alike because of their small size, which makes e-bikes a solution for short- to medium-distance trips. However, due to the limitation of the cycling environment and battery life, the trajectory points of shared e-bikes tend to exhibit discontinuities and non-uniformities, and it is difficult to obtain riding status information. A new method, multi-rule-constrained homomorphic linear clustering (MCHLC), is proposed to identify trips from the trajectories of shared e-bikes. From the perspective of the semantic trajectory, the new method constructs the stop/move model based on various criteria, such as speed, time duration, directional constraint, and contextual constraint, and identifies users' trips based on riding experience rules.

The main contributions of this paper are as follows:

(1) A new method is proposed to address sparse trajectory data. Shared e-bike trajectories are taken as an example to verify the performance of the method, which also enriches the research related to the shared e-bikes.

(2) The new method can effectively detect temporary stops resulting from specific purposes or behaviors (such as transporting children to school on the way to work), thereby revealing the details of residents' travel behaviors and providing scientific data for urban hotspot detection, hot route analysis, and urban functional zone sectorization.

The remainder of the paper is organized as follows. Section 2 describes the workflow of the new method, which is tested with the experimental data provided in Section 3. Section 4 discusses the role of the two factors (directional change constraint and contextual constraint) of the method and compares the new method with baseline algorithms. We draw conclusions and outline future work in Section 5.

2. Methodology

Certain related terms are first introduced to explain the new method.

Definition 1 (trajectory). A trajectory *T* is a series of discrete spatio-temporal points, each of which is a triple containing geographic coordinates and timestamps. A spatio-temporal point is expressed as $P_i = (x_i, y_i, t_i)$, i = 0, 1, ..., N, and $\forall 0 \le i < j \le N$, $t_i < t_j$.

As described in **Definition 1**, a trajectory essentially is a polyline consisting of a series of spatial points that are successive in terms of the timestamp. For example, in Figure 1, polyline T is a trajectory and is subject to the spatio-temporal point set $\{a, b, c, d, e, f\}$.



Figure 1. Illustration of a spatio-temporal trajectory [12].

Mao [36] summarized previous research on the mobility behavior of urban residents and defined a trip as follows. A trip is the movement of a resident from A to B using one or more transportation modes due to certain purposes. A reasonable trip is that the object moves at least a minimum distance in space and lasts for a minimum duration in time. Therefore, a trip corresponds to the movement along a trajectory, represented by a set of sequence points with a higher velocity.

A complete trip connects two consecutive stops: the origin and the destination. A stop along the trajectory indicates that a certain activity has been carried out at a certain location for a period of time. To better understand the behavior of residents, a temporary stay caused by a specific behavior or purpose is also considered a stop. For example, the temporary stay when dropping children off at school on the way to work is considered a stop, while the temporary stay caused by waiting at a traffic light is not a stop. From the perspective of data, a stop corresponds to a sequence of points with a velocity of zero or a very low velocity (such as below four kilometers per hour).

Definition 2 (trajectory element). *A trajectory element E, the basic linear unit of a trajectory, is the line connecting two adjacent points of the trajectory, as shown in Figure 1.*

In Figure 1, trajectory T contains five trajectory elements E. The attribute information of a trajectory element, such as the distance, average speed, and motion status, can be calculated by the endpoints. The length of a trajectory element is calculated by Equation (1), based on which the shortest distance between two points over the earth's surface can be calculated [37,38]. P_i and P_{i+1} are endpoints, and R is the average radius of the reference ellipsoid (WGS84), equal to 6371 km.

$$dis(P_{i}P_{i+1}) = R * \cos^{-1}(sin(P_{i}.lat*\frac{\pi}{180^{\circ}})sin(P_{i+1}.lat*\frac{\pi}{180^{\circ}}) + cos(P_{i}.lat*\frac{\pi}{180^{\circ}})cos(P_{i+1}.lat*\frac{\pi}{180^{\circ}})cos((P_{i+1}.lon - P_{i}.lon)*\frac{\pi}{180^{\circ}}))$$
(1)

The length of the trajectory element divided by the time interval is the average speed. The mathematical expression is shown in Equation (2).

$$\mathbf{E}_{i} \overset{=}{\vartheta} = \frac{\mathrm{dis}(\mathbf{P}_{i+1}, \mathbf{P}_{i})}{\mathbf{E} \cdot \Delta \mathbf{t}} = \frac{\mathrm{dis}(\mathbf{P}_{i+1}, \mathbf{P}_{i})}{\mathbf{P}_{i+1} \cdot \mathbf{t} - \mathbf{P}_{i} \cdot \mathbf{t}}$$
(2)

The motion status of a trajectory element can be obtained by comparing the average speed $E_i \cdot \vartheta$ to a given threshold ϑ_{thresh} , that is, if $E_i \cdot \vartheta$ is not greater than ϑ_{thresh} , the motion status is stopped, marked as "s", otherwise, the motion status is moving, marked as "m". The mathematical expression is as follows:

$$E_{i}.status = \begin{cases} s & E_{i}.\overline{\vartheta} \le \vartheta_{thresh} \\ m & E_{i}.\overline{\vartheta} > \vartheta_{thresh} \end{cases}$$
(3)

The new method is inspired by the velocity sequence linear clustering (VSLC) algorithm [39], which utilizes the semantic information of a trajectory to detect the stops due to the refueling behavior of taxies. Considering the trajectory element as the basic unit, the VSLC algorithm constructs a sequence of elements that have motion status according to Equation (3) and performs sequence merging of those elements with the same motion status to obtain sub-trajectory clusters of moving or stopping. The motion status of the sub-trajectory cluster is decided by the duration criterion. When the duration of the sub-trajectory cluster is shorter than the corresponding minimum duration threshold, the original motion status is pseudo and subject to revision. To build the stop/move model, homomorphic linear clustering is again performed after revising the motion status. Finally, the stops due to fueling are identified by analyzing the stop/move model with semantic information. The key to the VSLC algorithm is the correct identification of the motion status. The motion status of a temporary stay may be misidentified using only the duration criterion. Moreover, occasional behavior may divide a trip into multiple fragments, resulting in misidentification of the motion status by the duration criterion.

To solve such issues, two factors, directional change constraint and contextual semantic constraint, are introduced to determine the motion status of the sub-trajectory clusters. Trips are extracted by analyzing the stop/move model with the help of daily riding experiences. The new algorithm adopts the idea of clustering to extract trips from single trajectories, of which the methodological steps are shown in Figure 2.



Figure 2. Trip extraction workflow of the multi-rule-constrained homomorphic linear clustering (MCHLC) algorithm.

To better understand the principle of the MCHLC algorithm, all methodological steps are divided into three parts, which are implemented step by step in Figure 3. The first part is the formation of the original sub-trajectory clusters by sequentially clustering the trajectory elements with the same motion status. The motion status of a trajectory element is determined by comparing the speed and the given threshold. As shown in Figure 3a, one trajectory is composed of multiple trajectory elements marked "s" or "m", which expresses the motion status and is calculated by Equation (3). The sub-trajectory cluster marked "s" or "m" is formed by performing homomorphic linear clustering, as shown in Figure 3b.



Figure 3. Trip extraction of shared e-bikes by the MCHLC algorithm. (a) The motion status of each trajectory element is determined by the comparison of the speed and the given threshold. (b) The original clusters of moving or stopping are obtained by sequentially clustering the elements with the same status. (c) The status of the cluster in the red circle is misidentified using the duration criterion, which is correctly revised using the directional constraint, as shown in (d). (e) The trajectory is segmented into many segments, each of which is composed of clusters with the motion status in the form of " $m_1 - \cdots - s_n$ ". (f) The pseudo status is revised using the contextual constraint. (g) New clusters are obtained by performing homomorphic linear clustering again after status revision. (h) A trip is extracted using the rules based on daily riding experiences.

The second part is the construction of the stop/move model, which is the core of the MCHLC algorithm. In this part, the identification and revision of the motion status of the sub-trajectory clusters is the key to construct the stop/move model. According to the definition of a stop or a move, each motion status should last for a minimum duration over time. Thus, the duration criterion is first used to determine whether the identification of the sub-trajectory cluster is correct. That is, when the duration of a sub-trajectory cluster is not greater than the corresponding threshold, the current status of the sub-trajectory cluster is pseudo. The pseudo status indicates that the motion status may be misidentified, labeled as "F". Then, the directional constraint criterion is used to evaluate the sub-trajectory cluster with the pseudo motion status. For example, the pseudo status of the sub-trajectory cluster marked "s" is a true stop, whose direction change angle is near 180 degrees, which implies that the trips alternate. The direction change angle of the cluster is the difference between the heading angles of adjacent trajectory elements, calculated by Equation (4). The heading angle of a trajectory element is calculated by the geographic coordinates of the endpoints, following Equations (5) and (6). As Figure 3d shows, the misidentified status in the red circle of Figure 3c can be revised using the directional constraint. The latter shows that the directional constraint is useful to correctly identify the status of the temporary stay and can reveal detailed information of the trip.

$$dir_diff = |E_{i+1}.head - E_i.head|$$
(4)

$$E_{i}.head = \sin^{-1} \left(\frac{\sin(90 - P_{i+1}.lat) = \sin(P_{i+1}.log - P_{i}.lon)}{\sqrt{1 - \beta^{2}}} \right)$$
(5)

$$\beta = \cos\left((90 - P_{i+1}.lat) * \frac{\pi}{180}\right) * \cos\left((90 - P_i.lat) * \frac{\pi}{180}\right) + \sin\left((90 - P_{i+1}.lat) * \frac{\pi}{180}\right) * \sin\left((90 - P_i.lat) * \frac{\pi}{180}\right) * \cos\left((P_{i+1}.lon - P_i.lon) * \frac{\pi}{180}\right)$$
(6)

A whole trip may be divided into fragments by occasional behavior, such as waiting at a traffic light or avoiding pedestrians. Among the fragments, some motion status may be pseudo due to the short duration, resulting in misidentification of the trip. To address this issue, a contextual semantic constraint is introduced to determine and revise the motion status. Generally, a trip ends at a stop where the trips alternate. Therefore, a single trajectory is first segmented into multiple segments. Each segment is composed of multiple sub-trajectory clusters with a motion status in the form of " $m_1 - \cdots - s_n$ ", where the last sub-trajectory cluster marked " s_n " is a true stop. Certain sub-trajectory clusters in each segment may have the pseudo status, except for the status of the last sub-trajectory cluster. To eliminate misidentifications caused by the pseudo status, the contextual relationships of the clusters in a segment are analyzed according to actual riding experiences, which are summarized in Figure 4. When there is only one pseudo status in the segment, the pseudo status is regarded as noise caused by the statuses of adjacent segments, which should be revised. In Figure 4a, the pseudo status of the moving cluster may be caused by GPS signal drift, so the motion status of the cluster is revised from "m" to "s", denoting a stop. In Figure 4b, the pseudo status of the temporary stay caused by occasional behavior (such as waiting at a traffic light) is in the middle of the trip, thereby splitting the trip into fragments. Therefore, the status of the temporary stay should be revised to "m", denoting a move. If the occasional behavior occurs at the endpoint of a trip, there will be multiple pseudo statuses among the segments, as shown in Figure 4c, d. Then, the pseudo status of each cluster should be revised. Based on the contextual semantic constraint, the status of each cluster is re-identified and revised, and then homomorphic linear clustering is performed again to build the stop/move model.



Figure 4. The distribution of the pseudo state in different situations; (**a**) shows that the pseudo state of moving may be caused by GPS signal drift, (**b**–**d**) show the pseudo states caused by occasional behavior, which can occur in the middle, at the start, and at the end of a trip.

The third part is trip extraction from the stop/move model using the daily riding rules. According to daily riding experiences, a trip should satisfy the following rules:

(1) Distance rule: According to the definition, a trip should be at least a minimum distance in space. Here, the minimum distance is set as 500 m. The trip length is calculated from the distance between the points instead of the Euclidean distance between the endpoints of the trip.

(2) Average speed rule: Compared to shared bikes, the speed of shared e-bikes is higher. The report on sharing bicycles and urban development in 2017 stated that the average velocity of the fastest shared bicycles in cities is approximately 9.7 km/h [40], and here, we consider that the average velocity of a shared e-bike during a trip should be higher than 10 km/h.

(3) Shortest path rule: Generally, residents select the shortest travel path. If the distance between the endpoints of the trip is not greater than half the trip length, the trip should be filtered out.

(4) **Completeness:** A trip is usually connected by two consecutive stops. However, when a trip is at the endpoint of a trajectory, the trip may not be adjacent to the stop. In this situation, the endpoint of the trip can be decided by evaluating the speed. Considering that the speed gradually changes near the start or end point of the trip, when the speed at a point is not higher than 1.5 times the average speed of the trip, the point is considered an endpoint of the trip.

3. Experimental Data and Results

3.1. Experimental Data

The experimental data are the trajectory data of the shared e-bikes in Tengzhou city. As shown in Figure 5, the data cover the whole city center. Shared e-bikes are an emerging green travel mode, which is a solution to short- to medium-distance trips, especially in second- and third-tier cities. Compared to shared bikes, shared e-bikes with electric assistance have a distinct superiority in solving cycling barriers, such as longer trips and overcoming a challenging topography (hilly or dispersed cities) [41]. Moreover, shared e-bikes attract additional user groups who carry loads when traveling [42] or suffer from physical defects, which do not allow bicycle pedaling [43]. Fewer existing studies are related to shared e-bikes, most of which were based on survey data to study the performance of e-bikes [44], users' mobility behavior [45–47], and travel mode influencing factors [48]. Li and Dai [49] completed data cleaning of shared e-bike trajectories based on the speed and time interval rules for the first time, without considering the trajectory semantic information. To utilize the trajectory semantic information, a new method, MCHLC algorithm, is proposed to extract trips from the shared e-bike trajectories.

The shared e-bikes in Tengzhou city came from the BeeFly company and are named Mebikes, due to their bee-like appearance. Similar to dockless shared bikes, users can pay using a smart phone to pick up or return Mebikes freely, owing to the development of electronic fence technology. Users can be granted discounts on riding fees when e-bikes are returned to the electronic virtual stations set up using the electronic fence technology, which is an excellent solution to the issue of random parking.

The experimental data were acquired between May 19 and May 26, 2018, involving 516 Mebikes and 98,795 GPS trajectory points. The geographical coordinates are 117°07′14"–117°12′36"N and 35°02′08"–35°07′21"E. Each shared e-bike was equipped with an integrated GPS and communication module. The data were acquired from a specified internet address, where the GPS information is sent to every minute. As Figure 6 shows, the GPS records were stored as individual files by the key value of the vehicle ID. All the work was conducted using the Java programming language.



Figure 5. The coverage of the trajectory data; (**A**) shows the spatial distribution of the trajectory GPS points of a week in Tengzhou city, (**B**) is the enlarged display of the trajectory GPS points in the red box of (A), (**C**) shows the appearance of the shared e-bike, the Mebike, which is used in Tengzhou, (**D**) is the boundary of the shared e-bike usage shown by the app, and (**E**) is the electronic virtual station shown in the app.



Figure 6. Data storage format of the trajectory data.

As shown in Figure 6, each of the original GPS points includes the vehicle ID (the StationID), data acquisition time (timestamp), geo-location (latitude and longitude coordinates), and predicted mileage (anticipated mileage). It is not an easy task to extract trips from the raw data without any riding status-related attribute information. The GPS devices are set to collect data once a minute, but only 51.7% of the original data are recorded with a sampling interval of one minute. As shown in Figure 7, the original data have different sampling intervals. Among the data, 82.1% of the sampling intervals of the original data are shorter than 2 minutes, and 10.8% of the data are low-density sampling data, in which the sampling interval is longer than 2 minutes [50]. These statistics show that most of the data are uniform, but the data still have certain sparsity.



Figure 7. Distribution of the data sampling interval.

Data sparsity is caused by many factors, one of which is the difference in the travel time span of shared e-bikes. Figure 8 shows that 88.68% of the shared e-bikes are used on three to four days, and only 0.58% of the shared e-bikes are used on one day. The difference in time span causes data sparsity to a certain extent. Moreover, data sparsity is also related to the riding environment, and the battery can also result in discontinuous data. When riding between buildings or trees or along narrow alleys or when the battery runs out, the GPS signal will be weak or may be lost, resulting in discontinuous data or missing data. Such characteristics of shared e-bike trajectories indicate that a new method is needed to extract trips. Notably, the high utilization rate of shared e-bikes and the relatively uniform sampling rate both indicate that the data are usable and analyzable, although the data used in the study are sparse.



Figure 8. Utilization rate of the 516 shared e-bikes in one week (*x*-axis is the number of shared e-bikes corresponding to the utilization time; *y*-axis labels the utilization time in one week).

3.2. Experimental Results

As described in Section 2, the MCHLC algorithm depends on three key parameters, which are summarized in Table 1. Among these parameters, *Min_Move* and *Min_Stop* are the minimum duration time of a move and a stop, respectively, which are set to determine the corresponding motion status of a sequence. To reduce the misidentification of the stops caused by waiting for the traffic lights,

Min_Stop is set as 3 minutes to detect the true stops. According to the report on sharing bicycles and urban development in 2017, it takes approximately 3 minutes to complete a minimum trip (500 m). Compared to shared bicycles, residents usually choose shared e-bikes for a longer trip, so the parameter *Min_Move* is set as 4 minutes. Notably, a speed threshold is used to determine the original status of the trajectory element. Here, the speed threshold is set as the walking speed of 4 km/h. *Direction_Diff* is used to distinguish between a true stop and a temporary stay caused by special purposes, such as dropping children off at school on the way to work.

Parameter	Parameter Description			
Min_Move	The minimum duration for a normal move	4 min		
Min_Stop	The minimum duration for a normal stop	3 min		
Direction_Diff	The angle of the direction change between adjacent points	180°		

Table 1. The key parameters of the MCHLC algorithm.

Based on the default settings of the parameters listed in Table 1, the trajectories of shared e-bikes in Tengzhou were processed with the MCHLC algorithm. The invalid points beyond the study area were filtered out before utilizing the MCHLC algorithm. A total of 5962 trips were identified from the original data. The statistical results in Table 2 show that the shortest trip is approximately 833 m, while the longest trip length reaches up to 12.5 km. The average trip length is approximately 2.5 km, and the average riding duration is 10 minutes. To further examine the users' riding behaviors of shared e-bikes, the distribution of the identified trips is analyzed in the form of a bar chart (as shown in Figure 9). The majority of the trip lengths are within 1–4 km (accounting for more than 85% of all trips), of which 1–2 km accounts for the largest proportion (approximately 41.4%), followed by 2–3 km (nearly 30%), while 3–4 km accounts for the smallest proportion (only 14.8%). The bar chart shows that the trip length is mostly within 2–5 km when the residents of Tengzhou city choose shared e-bikes to travel. This result confirms that shared e-bikes are an excellent solution to short- to medium-distance travel in Tengzhou city.

Number of Trips	Μ	Minimum Travel Distance (meter)			aximum Distan (meter	Travel ce :)	Averag Dist (me	e Travel ance eter)	Aver	age Duration (minute)
5962		833.5	5938		12552.1	6	2466	5.534		10.36
	trip distance / kilometer	10.0-13.0 9.0-10.0 8.0-9.0 7.0-8.0 6.0-7.0 5.0-6.0 4.0-5.0 3.0-4.0 2.0-3.0 1.0-2.0 0.5-1.0	0.02% 0.05% 0.25% 0.45% 1.07% 2.73% 2.7%	6 7.06% 500	14.83% 1000	1500 umber of tr	29.32%	41.44%	3000	

Table 2. Statistical features of the extracted trips.

Figure 9. Distribution of the traveling distance for the extracted trips.

To verify the performance of the MCHLC algorithm, 50 shared e-bikes were sampled. A total of 486 trips were selected as reference data from the samples, which were obtained by manual interpretation in the ArcGIS Engine against the background of Amap and the Open Street Map (OSM)

road network. The MCHLC algorithm was evaluated by three data mining indexes (precision, recall, and F1-score), whose mathematical expressions are shown in Equations (7)–(9). Among these indexes, TP is the number of trips correctly detected, while FP is the number of trips incorrectly identified by the algorithm, and FN is the number of trips that the algorithm failed to identify. Table 3 indicates that all three indexes have a high value, higher than 92%. The results show that the MCHLC algorithm is reliable and suitable to identify trips from sparse trajectory data.

$$Precision = TP/(TP + FP) \times 100\%$$
(7)

$$Recall = TP/(TP + FN) \times 100\%$$
(8)

$$F1 - score = 2*Precision * Recall/(Precision + Recall) = 2TP/(2TP + FP + FN)$$
(9)

Index	Value	Index	Value
Referenced trips	489	Precision	95.62%
Calculated trips	479		
TP	458	Recall	93.66%
FP	21		
FN	31	F1-score	94.34%

Table 3. Evaluation of the MCHLC algorithm.

Note: TP: the trip that is successfully detected; FP: the trip is detected by the algorithm, but not true; FN: the trip, but failed to be detected.

4. Discussion

4.1. The Roles of the Directional Change Constraint and Contextual Constraint

Compared to the baseline VSLC algorithm, the two factors, directional constraint and contextual constraint, are mainly introduced in the new method to improve the performance. The directional constraint can effectively identify temporary stays, which may be caused by special travel purposes, even with a short duration. As shown in Figure 10, without the directional constraint, the GPS trajectory points in part A were misidentified as belonging to one trip, while the points were correctly identified as four trips considering the directional constraint, as shown in part B.

In Figure 10b, Trip 1_1 shows that the user departed from the New Star International Cinemas and arrived at the People's Harmony Bay Community (a residential area). The shared e-bike departed from the residential area 5 minutes later in the opposite direction, which indicated that a new trip occurred. Due to the missing data caused by the GPS signal being obscured by tall buildings, the stop cannot be identified using the speed criterion. However, considering the directional constraint, the stop can be identified, and the GPS points were identified as belonging to two trips, Trip 1_1 and Trip 1_2.

The shared e-bike departed from the People's Harmony Bay Community (a residential area) at 14:25 and passed by Beiwen Primary School at 14:34 before arriving at the True Love Shopping Mall at 14:47.

Trip 1_2 shows that the shared e-bike departed from the People's Harmony Bay Community (a residential area) at 14:25 and arrived at Beiwen Primary School at 14:34, while Trip 1_3 shows that the shared e-bike departed from Beiwen Primary School at 14:36 in the opposite direction and arrived at the True Love Shopping Mall at 14:47. We conjectured that a resident had dropped off their child at school on the way to the shopping mall, resulting in a two-minute stop at the school. Considering that this two-minute stay was caused by the resident's travel purpose, the stay was considered a stop as identified by the directional constraint, and two trips were distinguished.

A new trip, Trip 1_4, occurred between the True Love Shopping Mall and the Seven Degrees Network Club (a leisure area). The duration of the stop was only two minutes at the True Love Shopping Mall, where many people arrive and leave. We conjectured that someone had picked up the

shared e-bike soon after Trip 1_3 had ended and left the mall in the opposite direction. The short stop, misidentified by the speed and duration criteria, was correctly identified by the directional constraint.



Figure 10. The role of the direction change constraint; (**A**): the extracted trip without the direction change constraint, (**B**): the extracted trip with the direction change constraint.

It is noted that the directional factor used to identify the trip is based on the hypothesis that residents rarely turn 180 degrees during the trip without special reasons or purposes. If a person turns 180° in the middle of a journey, it means that some activity may have occurred. Even if the staying time is very short, it also can be identified based on the directional change factor, detecting the special purpose during a journey and revealing the detailed information of residents' travel behaviors. It is the advantage of our method over other methods. However, the directional change factor does not work if some activity with a short time has occurred without direction change. For example, in Figure 10, if the Primary School were at 90 degrees of the residential area, the trip 1_2 could not be identified due to the short stay and lack of direction change.

The pseudo stop caused by waiting at a traffic light can be revised by the contextual constraint. In Figure 11A, the trip extracted by the MCHLC algorithm is theoretically compatible with the shared e-bike departing from the Tengdu Community (a residential area) and arriving at a fitness club (a leisure area), in which the temporary stay (the enlarged figure in Figure 11B) shown with the red rectangle is caused by a traffic light. The temporary stay is misidentified as a stop by the method of Li and Dai, resulting in the whole trip being divided into two different trips (as shown in Figure 11C), which can be revised by the contextual constraint.



Figure 11. Temporary stop recognition based on the context constraints; (**A**) the trip with a temporary stop can be extracted correctly by the contextual constraint; (**B**) the temporary stop in the red rectangle is enlarged; (**C**) the trip was divided into two different trips based on the method proposed by Li and Dai's combined rules of the time interval and speed.

A temporary stay, such as waiting at a traffic light or avoiding pedestrians, may divide a trip into multiple fragments with different motion statuses. As shown in Figure 12A, the trip extracted by the MCHLC algorithm is theoretically compatible with the shared e-bike departing from the Sea Moon Community (a residential area) and arriving at the Garden Lane Community (another residential area), in which the pseudo stop caused by a traffic light occurs at the beginning of the trip (shown in the red rectangle), resulting in trip fragmentation. Thus, the beginning point of the trip is misidentified by the duration criterion of the VSLC algorithm, as shown in Figure 12B.



Figure 12. The role of the context constraints; (**A**) a trip extracted with the contextual constraint; (**B**) a trip extracted without the contextual constraint.

4.2. Comparison of the Different Methods

Two factors, directional constraint and contextual constraint, are introduced into the MCHLC algorithm, which is inspired by the VSLC algorithm. In addition, speed is one of the criteria used

to extract trips in the MCHLC algorithm. Thus, to test the performance of the MCHLC algorithm, two baseline methods are used here: CB-SMoT and VSLC. To understand the role of the directional constraint and contextual constraint better, the experimental result of the MCHLC algorithm is also compared with the results of the direction VSLC and semantic VSLC algorithms. Notably, the direction VSLC algorithm introduces the directional constraint into the VSLC algorithm, and the semantic VSLC algorithm introduces the contextual constraint into the VSLC algorithm.

The experimental results of the five algorithms were evaluated using three indexes (precision, recall, and F1-score), which are compared in Figure 13. The comparison shows that the new algorithm is better than the others, whose evaluation indexes are all above 93%.



Figure 13. Comparison of the different algorithms.

Compared to the baseline CB-SMoT algorithm, our method (MCHLC) has an obvious advantage for sparse trajectory data. As Figure 13 shows, the evaluation indexes of CB-SMoT are much lower than the indexes of other methods, especially the recall of CB-SMoT, which is just 55.14%. The results show that CB-SMoT is not suitable for trip extraction of sparse trajectory. CB-SMoT is an extension of DBSCAN, in which a core point is calculated by testing the average speed of adjacent points through two parameters *Eps* and *Mintime*. However, when only a few points meet the speed limit, it is difficult for CB-SMoT to discover the stops. For example, in Figure 14, there are two different stops. One is a stop with many points in a place, shown in Figure 14B, which can be discovered easily by CB-SMoT. The other is only one single point with a large time interval, shown in Figure 14C, which cannot be discovered by CB-SMoT. To address this issue, duration time is used to identify stops in our method. For the stop in Figure 14C, the motion between the two points is a stop, and the duration time is greater than the minimum time of a stop, so it can be discovered by VSLC and our method.



Figure 14. Different types of stops in the trips; (**A**) displays two trips with different stops; (**B**) is a stop with many points in a place; (**C**) shows another type of stop, which is only one point with a large time interval in a place.

Compared to the baseline VSLC algorithm, the performance of the MCHLC algorithm has been improved by approximately 10%, owing to the two additional factors, directional constraint, and contextual constraint. Although the directional and contextual constraints can individually improve the accuracy of the algorithm, the influence of the two additional factors is different. The results indicate that the contextual constraint improves the accuracy rate of the extracted trips by the algorithm, while the directional constraint results in an increasing number of trips being extracted successfully from the referenced data. The temporary stays caused by special purposes are mainly detected by the directional constraint, while the contextual constraint reduces the misidentification of the temporary stays, inevitably caused by the occasional behavior of residents that may result in trip fragmentation.

5. Conclusions

Shared e-bikes, an emerging transport mode, are favored by an increasing number of people. It is significant to study mobility behavior based on shared e-bike trajectories, as it can provide reasonable decisions for urban development. Moreover, it is worth studying the trip identification method, which is the basis of the trajectory data mining and trajectory analysis. In this paper, a new method of trip identification is proposed, which is tested by the trajectories of shared e-bikes in Tengzhou city. The conclusions drawn from the experimental results are as follows:

(1) The new algorithm, named the MCHLC algorithm, is reliable and suitable to identify trips from sparse trajectory data of shared e-bikes. Compared to the baseline VSLC method, the three evaluation indexes (precision, recall, and F1-score) have been increased by approximately 10% with the MCHLC algorithm, indicating that the new algorithm is better for shared e-bike trajectory data. Compared to the baseline CB-SMoT, the MCHLC algorithm has an obvious advantage of enabling sparse trajectory to identify stops. Due to data missing, a stop may be a single point. In this case, the stop can be identified by MCHLC with the duration time parameter, whereas it cannot be discovered by the CB-SMoT.

(2) The performance of the MCHLC algorithm is controlled by the directional and contextual constraints. The former can effectively identify the short stops caused for special purposes, which enriches detailed behavior identification. The latter utilizes semantic relationships to reduce misidentifications, especially the fragmentation caused by occasional behaviors (waiting at traffic lights or avoiding pedestrians).

Our new algorithm can also be applied to other forms of sparse trajectory data. Residents' activities at different scales can be identified by changing the parameter settings. In ongoing work, we will identify urban hotspots and hot travel routes on the basis of shared e-bikes trips, thereby providing a scientific decision-making basis for urban planning.

Author Contributions: Conceptualization, Xiaoqian Cheng.; methodology, Xiaoqian Cheng.; formal analysis, Xiaoqian Cheng and Weibing Du; software, Jianming Shen and Weibing Du.; validation, Zhaoxin Dai; formal analysis, Xiaoqian Cheng; investigation, Jianming Shen and Zhaoxin Dai; resources, Jianming Shen; data curation, Jianming Shen and Zhaoxin Dai; software: Weibing Duand Xiaoqian Cheng; writing—original draft preparation, Xiaoqian Cheng; writing—review and editing, Xiaoqian Chengand Chengming Li.; validation, Zhaoxin Dai; visualization, Weibing Du and Jianming Shen; supervision, Chengming Li; project administration, Chengming Li; funding acquisition, Chengming Li. and Weibing Du.

Funding: This research is jointly supported by the project of the National Natural Science Foundation of China, grant number 41871375 and 41601364; the project of the National Basic Surveying and Mapping project, grant number A1705; the project of Provincial Key Technologies R & D Program of Henan, grant number 172102210280.

Acknowledgments: The authors would like to thank the anonymous reviewers for their valuable comments and suggestions. The authors would also like to thank Kimberly Yasutis (AJE) for her English editing in enhancing the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

 Mou, N.; Zhang, H.; Chen, J.; Zhang, L.; Dai, H. A Review on the Application Research of Trajectory Data Mining in Urban Cities. J. Geo-Inf. Sci. 2015, 17, 1136–1142.

- 2. Zheng, Y. Trajectory Data Mining: An Overview. ACM Trans. Intell. Syst. Technol. 2015, 6, 1–41. [CrossRef]
- Jie, B.; He, T.; Ruan, S.; Li, Y.; Yu, Z. Planning Bike Lanes based on Sharing-Bikes' Trajectories. In Proceedings of the ACM Sigkdd International Conference on Knowledge Discovery & Data Mining, Halifax, NS, Canada, 13–17 August 2017.
- 4. Zhang, Y.; Mi, Z. Environmental benefits of bike sharing: A big data-based analysis. *Appl. Energy* **2018**, 220, 296–301. [CrossRef]
- 5. Kim, N.S.; Yook, D. Enhancing the economic benefit assessment of roadway investment through the application of value of time by trip length. *Transp. Policy* **2018**, *68*, 28–36. [CrossRef]
- 6. Yuan, N.J.; Zheng, Y.; Xie, X.; Wang, Y.; Xiong, H. Discovering Urban Functional Zones Using Latent Activity Trajectories. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 712–725. [CrossRef]
- 7. Zhang, X.; Li, W.; Zhang, F.; Liu, R.; Du, Z. Identifying Urban Functional Zones Using Public Bicycle Rental Records and Point-of-Interest Data. *ISPRS Int. J. Geo-Inf.* **2018**, *12*, 459. [CrossRef]
- 8. Ai, T.; Yang, W. The detection of transport land-use data using crowdsourcing taxi trajectory. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Prague, Czech, 12–19 July 2016; Volume XLI-B8, pp. 785–788.
- 9. Ying, Z.; Brussel, M.J.G.; Thomas, T.; Maarseveen, M.F.A.M.V. Mining bike-sharing travel behavior data: An investigation into trip chains and transition activities. *Comput. Environ. Urban Syst.* **2018**, *69*, 39–50.
- 10. Yu, L.; Wang, F.; Yu, X.; Song, G. Urban land uses and traffic 'source-sink areas': Evidence from GPS-enabled taxi data in Shanghai. *Landsc. Urban Plan.* **2012**, *106*, 73–87.
- 11. Yu, Y.; He, Z.; Song, Z.; Xi'an, F.; Wang, C. Investigation on structural and spatial characteristics of taxi trip trajectory network in Xi'an, China. *Phys. A Stat. Mech. Appl.* **2018**, *506*, 755–766.
- 12. Zhang, P.; Deng, M.; Shi, Y.; Zhao, L. Detecting hotspots of urban residents' behaviours based on spatio-temporal clustering techniques. *GeoJournal* 2017, *82*, 1–13. [CrossRef]
- 13. Zhao, P.; Qin, K.; Ye, X.; Wang, Y.; Chen, Y. A trajectory clustering approach based on decision graph and data field for detecting hotspots. *Int. J. Geogr. Inf. Syst.* **2016**, *31*, 27. [CrossRef]
- 14. Zheng, L.; Dong, X.; Xin, Z.; Tan, L.; Hang, L.; Li, C.; Liu, W. Spatial-temporal travel pattern mining using massive taxi trajectory data. *Phys. A Stat. Mech. Its Appl.* **2018**, *501*, 24–41. [CrossRef]
- 15. Luo, W.; Tan, H.; Lei, C.; Ni, L.M. Finding time period-based most frequent path in big trajectory data. In Proceedings of the ACM Sigmod International Conference on Management of Data (SIGMOD'13), New York, NY, USA, 22–27 June 2013.
- 16. Xia, Y.; Wen, H.P.; Zhang, X. Hot route analysis method based on trajectory clustering. *J. Chongqing Univ. Posts Telecommun.* **2011**, *23*, 602–606.
- 17. Yu, W. Discovering Frequent Movement Paths from Taxi Trajectory Data Using Spatially Embedded Networks and Association Rules. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 855–866. [CrossRef]
- 18. Wang, W.-J.; Li, X.-M.; Jiao, P.-F.; Xu, G.-Q.; Yuan, N.; Yu, W. Exploring Intracity Taxi Mobility during the Holidays for Location-Based Marketing. *Mob. Inf. Syst.* **2017**, 2017, 6310827. [CrossRef]
- 19. Xi, Z.; Guo, D. Urban event detection with big data of taxi OD trips: A time series decomposition approach. *Trans. GIS* **2017**, *21*, 560–574.
- Gao, Q.L.; Li, Q.Q.; Yue, Y.; Zhuang, Y.; Chen, Z.P.; Kong, H. Exploring changes in the spatial distribution of the low-to-moderate income group using transit smart card data. *Comput. Environ. Urban Syst.* 2018, 72, 68–77. [CrossRef]
- 21. Huang, J.; Levinson, D.; Wang, J.; Jin, H. Job-worker spatial dynamics in Beijing: Insights from Smart Card Data. *Cities* **2019**, *86*, 83–93. [CrossRef]
- 22. Long, Y.; Thill, J.C. Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing. *Comput. Environ. Urban Syst.* **2015**, *53*, 19–35. [CrossRef]
- Krumm, J.; Horvitz, E. Predestination: Inferring Destinations from Partial Trajectories. In Proceedings of the 8th International Conference on Ubiquitous Computing (UbiComp 2006), Orange County, CA, USA, 17–21 September 2006; pp. 243–260.
- 24. Du, J.; Aultman-Hall, L. Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues. *Transp. Res. Part A Policy Pract.* **2007**, *41*, 220–232. [CrossRef]

- 25. Li, Q.; Yu, Z.; Xing, X.; Chen, Y.; Liu, W.; Ma, W.Y. Mining user similarity based on location history. In Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS 2008), Irvine, CA, USA, 5–7 November 2008; pp. 298–307.
- 26. Jia, T.; Ji, Z. Understanding the Functionality of Human Activity Hotspots from Their Scaling Pattern Using Trajectory Data. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 341. [CrossRef]
- 27. Spaccapietra, S.; Parent, C.; Damiani, M.L.; Macedo, J.A.D.; Porto, F.; Vangenot, C. A conceptual view on trajectories. *Data Knowl. Eng.* **2008**, *65*, 126–146. [CrossRef]
- Alvares, L.O.; Bogorny, V.; Kuijpers, B.; Moelans, B.; Vaisman, A. A model for enriching trajectories with semantic geographical information. In Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems, Seattle, WA, USA, 7–9 November 2007; pp. 162–169.
- Palma, A.T.; Bogorny, V.; Kuijpers, B.; Alvares, L.O. A clustering-based approach for discovering interesting places in trajectories. In Proceedings of the 23rd Annual ACM Symposium on Applied Computing (SAC'08), Fortaleza, Ceara, Brazil, 16–20 March 2008; pp. 863–868.
- Rocha, J.A.M.R.; Times, V.C.; Oliveira, G.; Alvares, L.O.; Bogorny, V. DB-SMoT: A direction-based spatio-temporal clustering method. In Proceedings of the 5th IEEE International Conference on Intelligent Systems, London, UK, 7–9 July 2010; pp. 114–119.
- 31. Xiang, L.; Meng, G.; Tao, W. Extracting Stops from Noisy Trajectories: A Sequence Oriented Clustering Approach. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 29. [CrossRef]
- 32. Xiang, L.; Shao, X. Visualization and extraction of trajectory stops based on kernel-density. *Cehui Xuebao* **2016**, *45*, 1122–1131.
- Leiva, L.A.; Vidal, E. Warped K-means: An algorithm to cluster sequentially-distributed data. *Inf. Sci.* 2013, 237, 196–210. [CrossRef]
- 34. Júnior, A.S.; Moreno, B.N.; Times, V.C.; Matwin, S.; Anjos Formiga Cabral, L.d. GRASP-UTS: An algorithm for unsupervised trajectory segmentation. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 46–68. [CrossRef]
- Junior, A.S.; Time, V.C.; Renso, C.; Matwin, S.; Cabral, L.A. A semi-supervised approach for the semantic segmentation of trajectories. In Proceedings of the 19th IEEE International Conference on Mobile Data Management (MDM 2018), Aalborg, Denmark, 25–28 June 2018; pp. 145–154.
- 36. Ma, Y. Research on Residents Behavoir of Arrractive Areas and Spatio-temporal Feature Based on Taxi Trajectory Data—A Case of Kunshan City; Nanjing Norm University: Nanjing, China, 2014.
- 37. Chris, V. Calculate Distance, Bearing and More Between Latitude/Longitude Points. Available online: http://www.movable-type.co.uk/scripts/latlong.html (accessed on 22 October 2019).
- Rob, G. Calculate the Distance between Two Points in Your Web Apps. Available online: https://www.htmlgoodies.com/beyond/javascript/calculate-the-distance-between-two-points-inyour-web-apps.html (accessed on 22 October 2019).
- 39. Yang, W.; Tinghua, A.I. Refueling Stop Activity Detection and Gas Station Extraction Using Crowdsourcing Vehicle Trajectory Data. *Cehui Xuebao* **2017**, *46*, 918–927.
- 40. Mobike Global. The Mobike White Paper: Bike Sharing and the City. Available online: https://mobike.com/global/public/Mobike%20-%20White%20Paper%202017_EN.pdf (accessed on 18 October 2019).
- 41. Fishman, E.; Wei, H. Bikeshare: A Review of Recent Literature. *Urban Transp. China* 2016, *36*, 92–113. [CrossRef]
- 42. Rose, G. E-bikes and urban transportation: Emerging issues and unresolved questions. *Transportation* **2012**, 39, 81–96. [CrossRef]
- 43. Johnson, M.; Rose, G.; Johnson, M.; Rose, G. Extending life on the bike: Electric bike use by older Australians. *J. Transp. Health* **2015**, *2*, 276–283. [CrossRef]
- 44. Inagaki, T.; Mimura, Y.; Ando, R. An Analysis on Excursion Characteristics of Electric Assist Bicycles by Travel Behavioral Comparison Based on Trajectory Data. In Proceedings of the 15th International IEEE Conference on Intelligent Transportation Systems (ITSC 2012), Anchorage, AK, USA, 16–19 September 2012; pp. 433–437.
- 45. Cherry, C.; Cervero, R. Use characteristics and mode choice behavior of electric bike users in China. *Transp. Policy* **2007**, *14*, 247–257. [CrossRef]
- 46. Cherry, C.R.; Yang, H.; Jones, L.R.; Min, H. Dynamics of electric bike ownership and use in Kunming, China. *Transp. Policy* **2016**, *45*, 127–135. [CrossRef]

- 47. Lopez, A.J.; Astegiano, P.; Gautama, S.; Ochoa, D.; Tampère, C.M.J.; Beckx, C. Unveiling E-Bike Potential for Commuting Trips from GPS Traces. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 190. [CrossRef]
- 48. Campbell, A.A.; Cherry, C.R.; Ryerson, M.S.; Yang, X. Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. *Transp. Res. Part C Emerg. Technol.* **2016**, *67*, 399–414. [CrossRef]
- 49. Li, C.; Dai, Z.; Peng, W.; Shen, J. Green Travel Mode: Trajectory Data Cleansing Method for Shared Electric Bicycles. *Sustainability* **2019**, *11*, 1429. [CrossRef]
- 50. Yuan, J.; Zheng, Y.; Zhang, C.; Xie, X.; Sun, G. An Interactive-Voting Based Map Matching Algorithm. In Proceedings of the 11th IEEE International Conference on Mobile Data Management (MDM 2010), Kansas City, MO, USA, 23–26 May 2010.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).