International Journal of Geo-Information

MDPI

Article

Profiling Public Transit Passenger Mobility Using Adversarial Learning

Yicong Li [1], Tong Zhang [2,*], Xiaofei Lv [1], Yingxi Lu [3] and Wangshu Wang [4]

[1] Zhongnan Engineering Corporation Limited, Power China, Changsha 410014, China; lyc@msdi.cn (Y.L.)
[2] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[3] College of Architecture and Urban Planning, Hunan City University, Yiyang 413000, China
[4] Department of Geodesy and Geoinformation, TU Wien, A-1040 Vienna, Austria; wangshu.wang@tuwien.ac.at
* Correspondence: zhangt@whu.edu.cn

Abstract: It is important to capture passengers' public transit behavior and their mobility to create profiles, which are critical for analyzing human activities, understanding the social and economic structure of cities, improving public transportation, assisting urban planning, and promoting smart cities. In this paper, we develop a generative adversarial machine learning network to characterize the temporal and spatial mobility behavior of public transit passengers, based on massive smart card data and road network data. The Apriori algorithm is extended with spatio-temporal constraints to extract frequent transit mobility patterns of individual passengers based on a reconstructed personal trip dataset. This individual-level pattern information is used to construct personalized feature vectors. For regular and frequent public transit passengers, we identify similar transit mobility groups using spatio-temporal constraints to construct a group feature vector. We develop a generative adversarial network to embed public transit mobility of passengers. The proposed model's generator consists of an auto-encoder, which extracts a low-dimensional and compact representation of passenger behavior, and a pre-trained sub-generator containing generalization features of public transit passengers. Shenzhen City is taken as the study area in this paper, and experiments were carried out based on smart card data, road network data, and bus GPS data. Clustering analysis of embedding vector representation and estimation of the top K transit destinations were conducted, verifying that the proposed method can profile passenger transit mobility in a comprehensive and compact manner.

Keywords: transit mobility embedding; generative adversarial network; smart card data; public transit

Citation: Li, Y.; Zhang, T.; Lv, X.; Lu, Y.; Wang, W. Profiling Public Transit Passenger Mobility Using Adversarial Learning. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 338. https://doi.org/10.3390/ijgi12080338

Academic Editors: Christos Chalkias, Marinos Kavouras, Margarita Kokla, Mara Nikolaidou and Wolfgang Kainz

Received: 25 June 2023
Revised: 4 August 2023
Accepted: 11 August 2023
Published: 12 August 2023

1. Introduction

Modeling and profiling human mobility is essential to the understanding of travel mobility patterns of individual passengers and urban socio-economic structure [1]. Analyzing human transit behavior plays a vital role in capturing generic human activity patterns and the distribution structure of urban functional areas [2,3]. Public transit mobility manifests a collection of spatially–temporally varying travel patterns [4]. Therefore, it is crucial to effectively integrate spatio-temporal information to represent transit mobility patterns. Understanding patterns of transit behaviors is essential for enhancing smart transportation systems by providing more efficient and personalized services [5]. Prior human mobility research was primarily based on analytical models such as the gravity model [6] and radiation model [7]. However, these traditional analytical models, which aim to develop a universal model to capture human mobility, lack the capability to describe heterogeneous travel behavior [8]. Unlike traditional analytical models, data-driven models based on transit big data offer a novel approach to studying human mobility [9] and the relationships between different areas [10] within a city.

Public transport systems such as subways and buses carry an enormous number of daily passengers in the city, including not only commuting activities but also tourist activities [11]. A large amount of smart card data (SCD), generated by smart public transport systems, has become an important data source for studying human transit mobility and further profiling passenger transit behavior from a data-driven perspective [12]. Additionally, through comparative studies, it is possible to obtain insights into the explainability of artificial intelligence concerning human mobility [13]. Unlike traditional statistical models that describe people's transit patterns from a group perspective, data-driven models usually focus on individual transit mobility. Typical data-driven models such as Recurrent Neural Network (RNN) and Word2Vec have achieved favorable results in practical application because these models can capture the temporal correlation of individual transit activity [14–17]. However, these models still have room for improvement, such as: (1) excessive attention is given to models with large data samples, while the modeling of spatio-temporal characteristics of transit is insufficient; (2) most studies did not explicitly account for complete trips, thereby failing to capture comprehensive semantic travelling information.

In recent years, the generative adversarial network (GAN) has become one of the most promising unsupervised learning models [18]. Compared with typical discriminative models, GANs have the following advantages: (1) as a generative model, they achieve good practical benefits in data imputation, data generation, and simulation of specific data distributions; (2) due to their high scalability, it is convenient to incorporate prior knowledge into the network to improve stability and accuracy [19]. This paper innovatively proposes to capture spatio-temporal characteristics of public transit passengers using the framework of the GAN by constructing high-level compact vector representations that describe passengers' transit mobility.

The novelties of this study can be summarized as follows:

1. We propose to profile public transit mobility considering both the personalized and group characteristics. We extend Apriori algorithm [20] with spatio-temporal constraints to extract passenger's frequent transit mobility, which representing personalized characteristics. And the group characteristics are represented by identifying similar transit group mobility. Additionally, because of the sparsity of public transit trajectory [21], we design a unified grid-based method to construct transit mobility vectors.
2. We propose a novel data-driven method for profiling public transit mobility using an adversarial learning network that integrates personal and group characteristics. The generator of the network consists of a pre-trained sub-generator and an auto-encoder, both of which are composed of multiple GRU layers and fully connected layers. We pre-train sub-generator using similar transit group vectors as labels to add group characteristics into the network. The discriminator, on the other hand, is composed of multiple fully connected layers. Within the framework of GAN, frequent transit mobility vectors are used as the real value to be jointly trained with the fake value of the generator. Through this adversarial process, a specific public transportation passenger's transit mobility embedding can be obtained;
3. The proposed method was evaluated with real SCD in Shenzhen, China. The experimental results show that the proposed approach has the applicability to characterize transit mobility and assist in understanding passengers' transit behavior and mobility patterns.

2. Related Work

2.1. Transit Mobility Pattern Mining

The classical model for studying group transit mobility patterns traces back to the gravity model proposed in the 1940s [6]. This model suggests that the intensity of human transit mobility between cities is proportional to the product of the population of the two cities and inversely proportional to the distance between them. Barabasi et al.'s seminal article [22] clearly revealed the temporal deviation of human behavior from the Poisson

process. Since then, there have been numerous studies on human activity patterns. Ref. [23] analyzed human behavior by studying the flow of banknotes and found evidence for the spatial scaling law of human behavior. Ref. [24] discovered the high regularity of human trajectory and the high probability of returning to specific positions. Ref. [25] found an upper-bound potential predictability rate of individual passengers of 93%. Ref. [26] proposed a nonparametric model based on Gaussian process regression to model passenger trajectories. Furthermore, how to extract significant human mobility patterns also attracted attention. Through a Markov chain [27], 17 significant mobility patterns were extracted by analyzing daily activity trajectory data. Ref. [28] proposed a complete trajectory pattern mining framework to extract frequent patterns of individual trips based on the heuristic detection approach. These studies offer enlightenment for analyzing transit mobility patterns by establishing classical statistical models, investigate transit data flows, and aggregating passenger trajectories. However, the aforementioned studies, relatively, lack spatio-temporal modeling, and they are unable to capture more transit characteristics due to the lack of substantial transit data.

2.2. Data-Driven Transit Mobility Analysis

Smart Card Data (SCD) recorded by public transportation systems contain rich information about people's activities within the city, providing a solid data foundation for understanding and studying human activities. Some studies systematically analyzed the application value of SCD and demonstrated how SCD can be used to generate and reconstruct passenger travel trajectories [29,30]. And several research papers have applied SCD to study people's travel activities and explore the relationship between human and cities. SCD were used to analyze the flow of people between areas and to characterize regional features, leading to the extraction of compact vector expression for regional representation [31]. Ref. [12] used one month of Beijing SCD to develop a data mining model for extracting and modeling passenger commuting patterns, and through comparison analysis of the work and residence locations of different commuters, verifying the imbalance between work and residence in the Beijing region. Some SCD cluster algorithms, such as [32], were applied to mine passenger transit pattern and detect daily transit behavior differences by augmenting classical algorithms. These studies indicate that SCD has significant value in areas such as analyzing passenger behavior, optimizing road networks, and adjusting public facilities.

In recent years, many researchers have focused on profiling public transit mobility from a data-driven perspective using deep learning techniques. These methods aim to capture the spatio-temporal characteristics of transit mobility and explore individual and group behavior patterns. By applying deep learning models such as Seq2Seq [33], Stacked-LSTM [17], Variational Auto-Encoder [34,35], Auto-Encoder [36], Word2Vec [37] and bidirectional self-attention network [38], these studies embedded passenger transit mobility and clustered the vector representations to explore mobility patterns, infer trip purpose, predict passenger activity in the next moment and solve trajectory matching problems. As human mobility is strongly spatio-temporal, some studies tended to use mobility graph to capture spatial characteristic. Ref. [39] established a probability pattern graph to express drivers' driving behavior and applied the representation results to the detection of high-risk sections of urban traffic accidents. Ref. [40] divided a passenger's trips into passenger, transit mode, and OD triplet structure to construct a behavior graph with heterogeneous edges. By jointly embedding passenger, transit mode, and OD, the study can be applied to the recommendation of the passenger's transit mode. Ref. [41] profiled individual mobility using location-based graphs and clustered based on different hyper parameter configurations.

From the perspective of generating models, some studies applied generative adversarial network [16,42] and GPT-2 [43] to embed human transit mobility. Ref. [16] established a probability pattern graph using transit behavior based on the sub-structure of individual behavior patterns and embedded passenger transit mobility by generating an adversarial network, while Ref. [42] proposed a semi-supervised learning model to solve the problem

of user trajectory matching, in which adversarial networks were used to regularize the data distribution of user trajectory. In summary, these studies used various deep learning models to capture spatio-temporal characteristics and explored individual and group behavior patterns in public transit mobility. They have addressed various problems, such as trajectory matching, trip purpose inference, behavioral preference exploration, detection of high-risk areas of urban traffic accidents, recommendation of transit modes, and prediction of passenger activities. However, there are still few methods to profile public transit mobility, and existing models seldom take into account the regularity and sparsity of public transit trips.

3. Methodology

3.1. Problem Formulation

In this study, we reconstruct public transit passenger trips using multi-source data, including SCD, GPS data of buses, and road network data (details will be further described in Section 4.1). Given a duration (5-day weekday or 2-day weekends), a passenger P_i has generated m trips, which constitute a passenger's personal trip dataset in the order of trip start time. We convert this personal trip dataset into a vector representation $P_i = \{Trip_1, Trip_2 \ldots, Trip_m\} \in R^{n \times m}$, where n is the original dimension of each trip, and m will vary based on different passengers and the set duration. To solve the variable length problem and make the vector representation more compact and better able to capture high-level semantic information, we need to find a mapping function $P_i' = \varphi(P_i) \in R^N$, where N is a constant value, to embed the original personal trip features into a low-dimensional R^N space.

3.2. Methodology Overview

From the perspective of individual public transit passengers, we propose a novel data-driven method to extract the low-dimensional and compact vector representation of public transit behavior using GAN, based on the frequency, regularity, and sparsity of public transit trips. The proposed method consists of the following four steps (as shown in Figure 1): (1) reconstructing passenger trips and building a personal trip dataset using transit network data, smart card data and GPS data; (2) representing personal trips as raw features and constructing grid-based transit mobility vector; (3) transit mobility embedding using a generative adversarial model that contains a pre-trained sub-generator and an auto-encoder; (4) application of the embedding results, including transit spatio-temporal pattern clustering analysis, transit mobility visual analysis and estimating top K transit destinations. Experiments were performed on a real-world dataset collected in Shenzhen City, China, indicating that the proposed method can effectively capture the high-level semantic information of public transit behavior and provide a more compact and low-dimensional representation to enhance the understanding of passenger transit behavior and mobility patterns.

3.3. Constructing Transit Mobility Vector

3.3.1. Representing Personal Trips as Raw Features

Suppose that a trip of a passenger starts from location o at time t_o, and arrives at location d at time t_d, we use a quadruple defined in Equation (1) to describe the transit trip.

$$Trip_j = (o, t_o, d, t_d) \tag{1}$$

In this paper, we consider trip feature engineering from three main aspects: time, space, and semantics, and the mapped dimensions are 4, 4, and 3, relatively. Firstly, we map t_o and t_d onto a two-dimensional circle respectively, using the idea of projection to reflect the continuity and periodicity of time. This enables the time feature to be represented by the (x, y) coordinates on the circle, which is defined in Equations (2) and (3), which *Time* indicates departure time or arrival time and T (T = 24) indicates the time frame is a full 24 h

day. Secondly, we use the origin and destination positions in the geographical coordinate system as the space feature. Thirdly, we extract trip duration and the number of passed stops to encode time distance and space distance, respectively. Additionally, we extract the number of transfers to describe the passenger's transfer behavior. Through these three parts, we can construct the raw transit mobility vector as $Trip_j \in R^n (n = 11)$. To accelerate the convergence of the deep learning model and keep the data at the same scale, we normalize each column of features and keep the data in the standard normal distribution, as defined Equation (4), which μ and σ represents the mean and the variance respectively. The normalization improves the performance of the model and ensures consistency of the data. However, for the personal trip dataset, this raw transit mobility vector representation can be high-dimensional and unable to capture the correlation among trips, making it challenging to be practically used. To address this issue, we propose to utilize this raw vector as the input of the GAN, which generates a more compact and low-dimensional transit mobility vector.

$$x = \cos\left(\frac{2\pi \times Time}{T}\right) \tag{2}$$

$$y = \sin\left(\frac{2\pi \times Time}{T}\right) \tag{3}$$

$$x' = \frac{x - \mu}{\sigma} \tag{4}$$

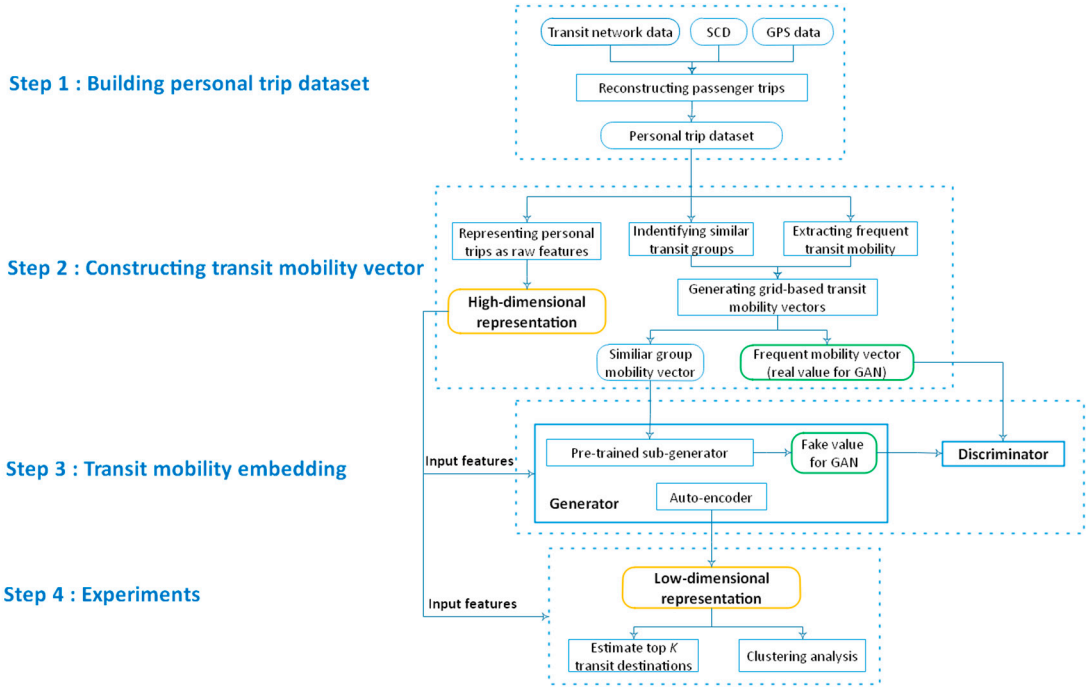

Figure 1. Workflow of the proposed approach.

3.3.2. Extracting Frequent Transit Mobility

Public transit mobility exhibits several key characteristics that distinguish it from other modes of transportation. These include: (1) regularity and frequent transit mobility: public transit users tend to exhibit regular commuting patterns, often traveling to and from work or school at similar times each day or week. This results in a high frequency of certain transit behaviors; (2) co-occurrence: transit trajectories of different passengers tend to occur together in time and space, meaning that a user's trip may be influenced by other passengers and their transit behavior; (3) spatio-temporal constraints: transit mobility is

subject to certain spatio-temporal constraints, such as the schedule of transit routes or the layout of the transit network.

Due to the frequency and regularity of public transit behavior, frequent trips can indicate personalized characteristics for individual transit patterns. When someone consistently takes specific trips on public transit, it suggests that those trips are tailored to their specific needs or preferences. It is natural to apply association analysis algorithms to mine frequent patterns in the personal trip dataset. In this study, we extend the Apriori algorithm [20], the classical association analysis algorithm, to extract frequent transit mobility by adding spatio-temporal constraints to the calculation of support. That is, the frequent transit mobility indicates the most frequent transit sequences set of this passenger (e.g., {[O_1, D_1], [O_2, D_2, O_3, D_3]}). If a passenger P_i consists of m trips (sorted in ascending order according to start time), and his personal trip dataset can be expressed in Equation (5), in which O_i, D_i, and d_i means the origin, destination, and transit time cost of the i-th trip, respectively. Trip datasets for all passengers can be expressed as the set $\{P_1, P_2 \ldots, P_m\}$. Referring to the definitions of association analysis, trip datasets for all passengers can be treated as original input dataset, P_i can be considered as an event, and O_i and D_i can be considered as the items of the event.

$$P_i = \left\{ \underbrace{O_1 \overset{d_1}{\rightarrow} D_1}_{trip_1} \overset{d_2}{\rightarrow} \underbrace{O_2 \overset{d_3}{\rightarrow} D_2}_{trip_2} \rightarrow \ldots \rightarrow \underbrace{O_m \overset{d_{2m-1}}{\rightarrow} D_m}_{trip_m} \right\} \tag{5}$$

The proposed algorithm for mining frequent patterns in passenger transit behavior consists of three steps (as shown in Algorithm 1). Firstly, as shown in lines 1–6, a triplet set $Q = (O, D, aveTime)$ is defined to represent the average time cost from O to D. The whole dataset is traversed to obtain the candidate set R_1 and set Q. Then the frequent set L_1 is obtained by applying a support threshold to the set R_1. Secondly, as shown in lines 7–12, a loop for extending the candidate set R_k and the frequent set L_k is executed (where k represents the length of the item). A set L_{all} is defined to store all the extracted frequent patterns if L_k satisfies the minimum length restriction. Then we generate R_{k+1} by connecting R_k with itself and pruning with spatio-temporal constraints. According to the set Q, we retain the candidates whose time costs satisfy the range of $[aveDur * (1 - \sigma), aveDur * (1 + \sigma)]$ with the corresponding O and D. Similarly, we can generate L_{k+1} by R_{k+1} and the support. Finally, we reconstruct the personal trip dataset using frequent mobility pattern set L_{all}. For a personal trip dataset, we determine whether its transit mobility contains frequent patterns. If the count of frequent patterns is greater than threshold δ, these frequent patterns are kept and the rest of trips are abandoned. If the count of frequent patterns is less than the threshold δ, these kinds of passengers can be considered as random or having strong personalized characteristics. In this case, these passengers lack frequent mobility patterns and all their trips are retained.

Due to the frequency and regularity of public transportation, passenger's frequent transit mobility is a key indicator of their transit behavior. We propose a grid-based method (as described in Section 3.3.4) to transform this mobility into a vector form, which can then be used as the real value to train a GAN.

Algorithm 1: Extracting Frequent Mobility Pattern.

Input: Trip datasets of all passengers in M days: totalTripDatasets = {P_1, P_2, P_3, . . ., P_{m-1}, P_m}; trip time threshold: σ; support threshold: θ; minimum frequent pattern length: γ; frequent pattern count threshold: δ
Output: Collection of frequent trips for all passenger in M days: $ETrips$ = {ET_1, ET_2, ET_3,. . ., ET_{n-1}, ET_n}

```
Initialize Q← ∅, L_all← ∅, k←1, j←1
for each tripDataset in totalTripDatasets do//Initializing data
  R_j.append(tripDataset.places)
  Q←UpdateQ(tripDataset)
end for
L_j←ExtractFrequentPattern(R_j, θ)//Extracting L_1 pattern
while len(L_k) > γ do//Loop for expanding patterns
  L_all.append(L_k)
  k←k + 1
  R_k←ExpandRSet(L_j, L_k, Q, σ)
  L_k←ExtractFrequentPattern(R_k, θ)
End while
ETrips = ReconstructTripDataset(L_all, totalTripDatasets)
```

3.3.3. Identifying Similar Transit Groups

Public transit passengers are often regular commuters who have daily traveling habits. Therefore, identifying similar transit groups to which these passengers belong can provide valuable insights into their group characteristics. That is, we aim to identify a similar transit group which a passenger belongs to. And we propose to use these trips of the transit group to represent the group characteristics for individual transit. We propose three soft constraints to extract similar transit groups for a passenger: (1) soft space constraint: the trips should be in close proximity, such as within 1 km, for both the origin and destination; (2) soft time constraint: the departure time, arrival time, and trip duration should be within a certain range, such as 1 h; (3) soft semantic constraint: the transit frequency, numbers of transfer, and whether a passenger is a commuter should exhibit some degree of similarity. In this study, by means of the grid-based method (as described in Section 3.3.4), we transform passenger's similar transit group mobility into the vector form, which can be used to pre-train the sub-generator of the GAN.

3.3.4. Generating Grid-Based Transit Mobility Vectors

Given the sparsity of public transit trajectories, we need a unified data structure to construct transit mobility vectors. We propose to use a regular grid to profile public transit mobility, and recursive quad-division of the study area is performed based on the popularity of transit trips, as shown in Algorithm 2. Firstly, we determine the study area scope by defining its lower left coordinates and upper right coordinates. These coordinates define the minimum bounding rectangle of the study area. Secondly, we quad-divide the study area into smaller grids based on certain criteria, such as meeting a maximum popularity limit and a minimum length limit. The popularity of a grid can be calculated by the number of trips that occur within it. After dividing the study area into grids, we evaluate each grid to determine if it satisfies the partition condition. If it meets the criteria, we further split it into smaller grids. This process is performed recursively until the partition condition is no longer satisfied. Finally, we collect all the split grids to form the result set, which represents the subdivided grids that meet the specified criteria within the study area.

Algorithm 2: Constructing grids based on trip popularity.

Input: Area scope: $\boldsymbol{G}$ = [lowerLeftLng, lowerLeftLat, upperRightLng, upperRightLat]; minimum grid length: γ; maximum flow: $\boldsymbol{\theta}$
Output: A regular grid: $\boldsymbol{FG}$={$\boldsymbol{fgrid_1}$, $\boldsymbol{fgrid_2}$,. . ., $\boldsymbol{fgrid_n}$}
1: Initialize $\boldsymbol{FG}\leftarrow\emptyset$, $\boldsymbol{SG}\leftarrow\boldsymbol{G}$//SG means the collection of split grids
2: for each $\boldsymbol{sg}$ in $\boldsymbol{SG}$ do
3: CalculateTransitPopularity(sg)//Calculating transit popularity for each grid
4: if $\boldsymbol{sg.flow} > \boldsymbol{\theta}$ and $\boldsymbol{sg.width} > \gamma$ and $\boldsymbol{sg.height} > \gamma$ then
5: $\boldsymbol{gTmp}$ = SplitGrid($\boldsymbol{sg}$)//Quad-dividing grids
6: FilterByQuatree($\boldsymbol{gTmp}$)//Recursive function
7: else then
8: $\boldsymbol{FG}$.append($\boldsymbol{sg}$)//Appending grids that meet the requirements
9: end if
10: end for

We represent a trip as a pair of associated grid units, based on its origin and destination. We construct a two-dimensional matrix for calculating trip probability according to the count of split grid. Then passenger's transit mobility vector can be represented as the joint trip probability of each grid cell. To be specific, we can obtain k grids by applying Algorithm 2 to the study region. A two-dimensional matrix $M \in R^{k*k}$ is established with each element $m(i,j)$ in M representing the count of trip which starts from i and arrives at j. It can be seen that the i_th row of M represents all the trips starting from i, while the j_th column represents all the trips arriving at j. For each row in M, we can calculate the probability which starts from i and arrives at j, namely $m'(i,j) = \frac{m(i,j)}{\sum_j m(i,j)}$. We utilize the maximum of these probabilities to represent the mobility of i_th row, then a vector of length k is obtained to express the departure feature of the passenger, namely $mm'_i = \max(m'(i,1),\ldots,m'(i,k))$. Similarly, at the view of column, we can also obtain a vector of length k to express the arrival feature of this passenger, namely $mm'_j = \max(m'(1,j),\ldots,m'(k,j))$. These two vectors above are stacked to obtain a vector with a length of 2*k, which is the vector representation of passenger transit mobility (Figure 2). By this grid-based method, we can vectorize the frequent mobility (as described in Section 3.3.2) and similar group mobility (as described in Section 3.3.3).

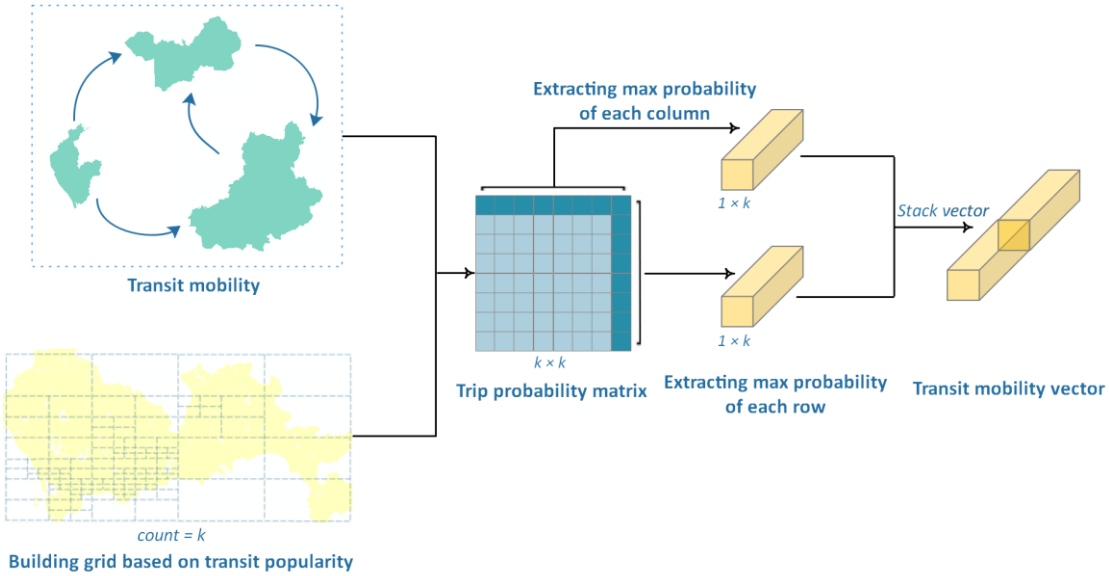

Figure 2. Constructing transit mobility vector.

3.4. Transit Mobility Embedding

3.4.1. Pre-Training Sub-Generator

Due to the regularity of public transit passengers, passengers with similar spatiotemporal characteristic can form transit similarity groups. The group characteristics of public transit passenger are reflected by these similar transit groups. To model passenger transit mobility, it is important to consider the individual's group characteristics. Using a vanilla GAN may result in unstable training and performance degradation. To address these issues, we propose using a pre-trained sub-generator (shown in Figure 3) to enhance the stability of the training process and comprehensively profile group characteristics of individual passengers.

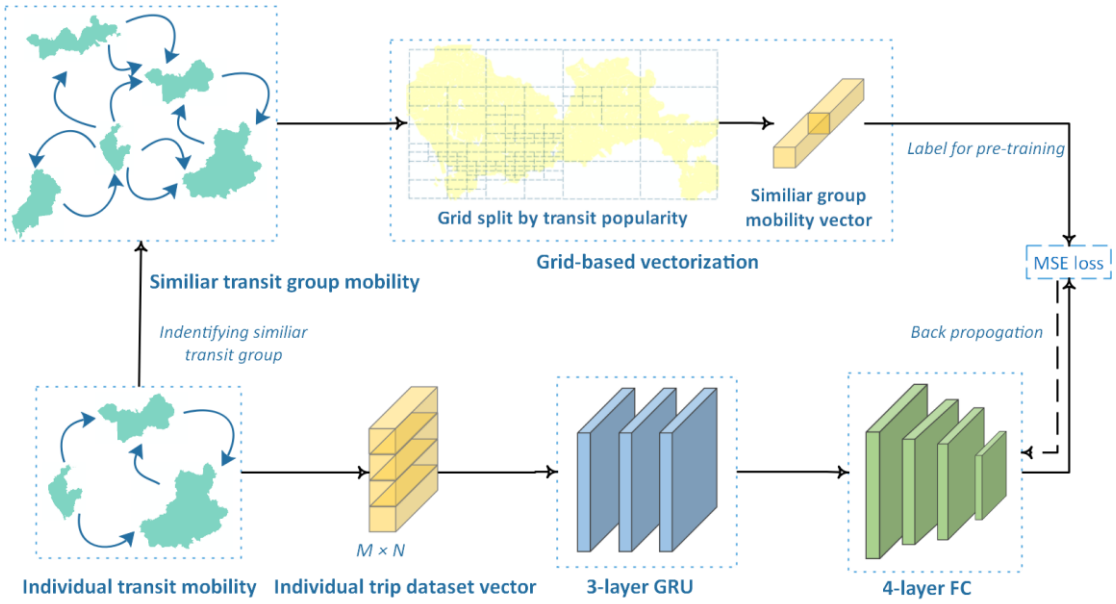

Figure 3. Pre-trained sub-generator.

After performing feature engineering (as described in Section 3.3.1), we can represent personal trip dataset as a matrix of $M \times N$ dimension, where M represents the count of trips and N represents the count of extracted trip features. Since adjacent trips may have strong temporal correlations, we use a multi-layer GRU (Gated Recurrent Unit) to capture the temporal correlation of the trips. The purpose of using a multi-layer GRU is to enable the network to further capture trip features and strengthen the deep learning model's ability to obtain temporal correlations among trip. Following the three-layer GRU, four fully connected (FC) layers, with sizes of 44, 80, 160, and $2 \times k$ (where k is the count of grids), reshape the dimension to fit the target label which is the similar transit group mobility vector (as described in Sections 3.3.3 and 3.3.4).

3.4.2. Embedding Mobility Using Adversarial Learning

Based on the two-player game idea introduced by [18], we propose a novel GAN for modeling passenger transit mobility. The input of the model is individual trip dataset vector, which obtained in Section 3.3.1. The generator comprises of an auto-encoder and a pre-trained sub-generator. The auto-encoder compresses and embeds the trips into a low-dimensional vector representation using GRU and FC structures, capturing the temporal dependencies between adjacent trips. After the encoding, the dimension of the trip vector remains unchanged, and the result is passed forward into the pre-trained sub-generator, which generates fake values to mislead the discriminator. The real value of the model is represented as the vector of individual frequent transit mobility, extracted in Section 3.3.2 and vectorized in Section 3.3.4. The discriminator is constructed with a 4-layer FC layer,

and a sigmoid function maps the output to a probability within (0, 1). The closer the probability is to 1, the higher the probability that the sample produced by generator is considered to be real, and vice versa. The cross-entropy function is used to calculate the loss and backpropagation is performed accordingly (Figure 4).

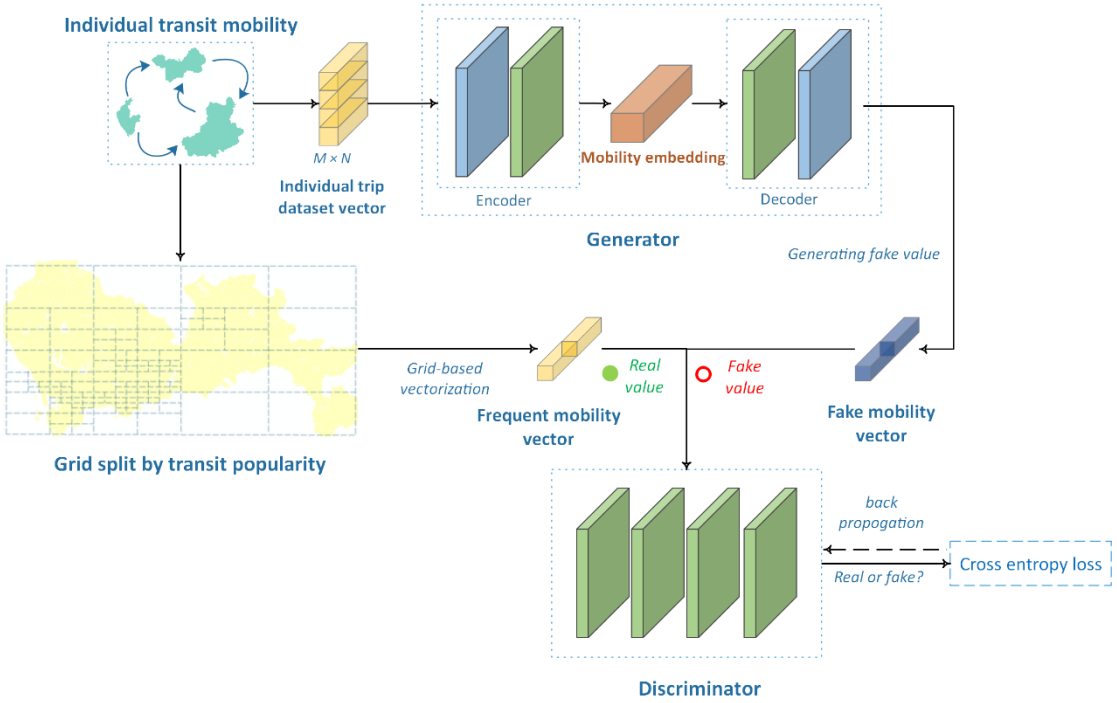

Figure 4. GAN-based mobility embedding.

After training the GAN model to convergence, we retain its network parameters and use the encoder of the auto-encoder to obtain the embedding of the personal trip dataset (Figure 5). This embedding is a compact and low-dimensional vector with comprehensive semantic information, which contains the personalized and group characteristics of a passenger's transit mobility. This vector representation plays a crucial role in analyzing public transit behavior, profiling public transit mobility, and supporting further studies of human transit patterns. In this paper, we utilize this compact vector representation to conduct experiments, such as visual analysis, cluster analysis, and estimating top K transit destinations.

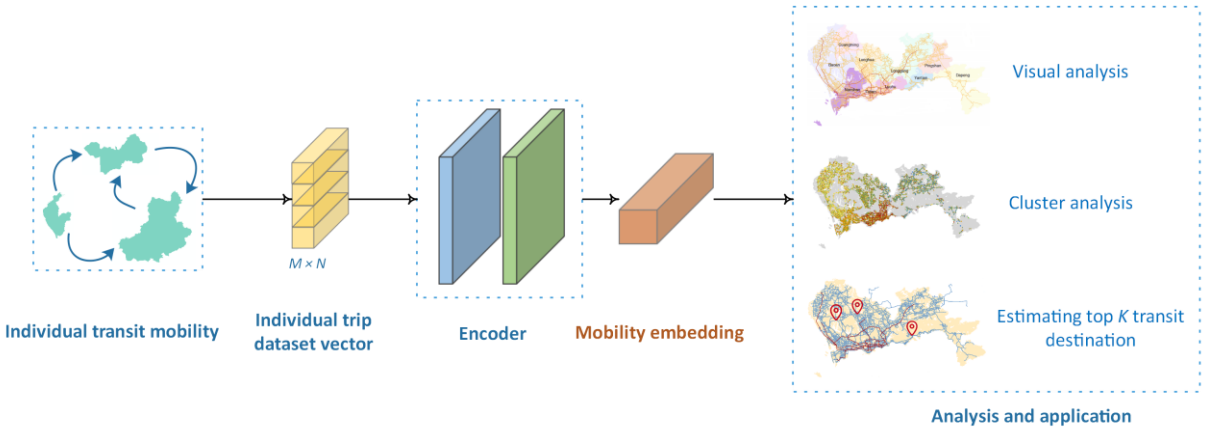

Figure 5. Generating mobility embedding.

4. Results and Discussion

4.1. Study Area and Data

The study area is located in Shenzhen City, China. Shenzhen is a populous city with approximately 17.5 million residents and encompasses an area of approximately 2000 km^2. The city boasts an extensive public transportation network. We collected SCD which includes data from both the bus and subway systems of Shenzhen's public transportation. Additionally, we obtained GPS data from buses and transportation network data of the city. The public transportation system in Shenzhen consists of 8 subway lines, 199 subway stations, 808 bus routes, and 6226 bus stops, as detailed in Figure 6. This comprehensive system provides a range of options for commuters and travelers navigating the city.

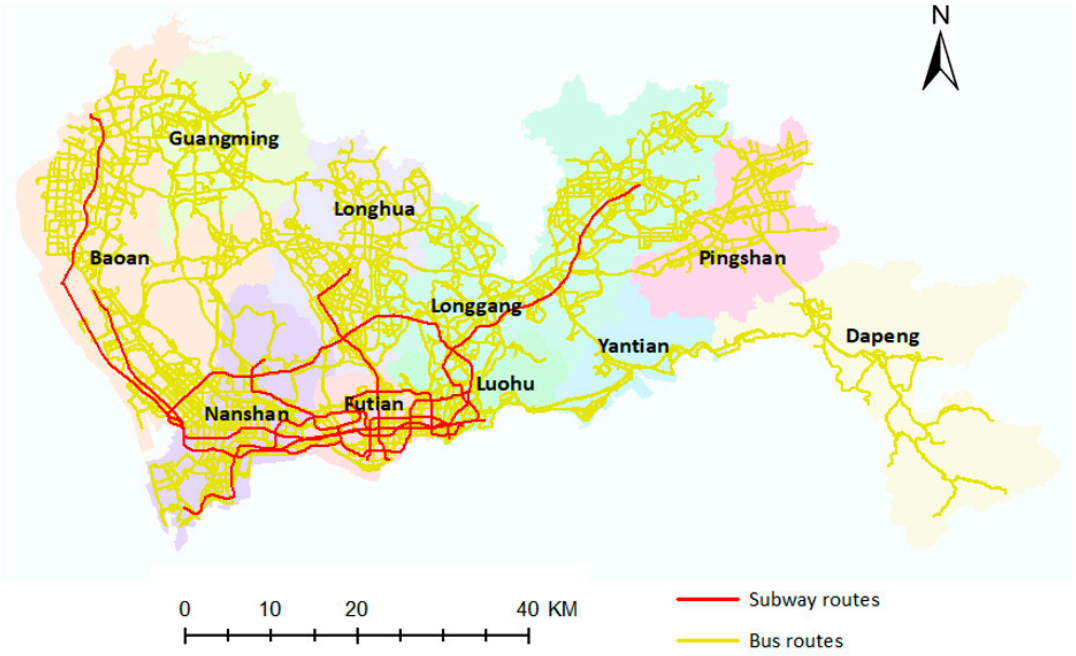

Figure 6. Shenzhen City and its public transit network.

We used one week of data for the evaluation, where the amount of SCD is about 7.8 G. On weekdays, the amount of SCD is about 1.2 G per day and the number of records is about 8.6 million. On weekends, the amount of daily SCD volume is about 0.9 G and the number of records is about 7.2 million. The details of SCD can be seen in Table 1, where the card type field "21" represents subway entry, "22" represents subway exit, and "31" represents bus boarding. And the swiping times vary between different bus lines and subway stops.

Table 1. Smart card data examples.

ID	Card Type	Swiping Time	Bus Companies (Subway Lines)	Bus Lines (Subway Stops)	Bus Number (Subway Gate)
1000051	22	17 April 2017 14:31:53	Subway 2	Yannan Station	OGT-122
200044	21	17 April 2017 09:01:08	Subway 6	Buxin Station	OGT-242
2000341	31	7 April 2017 17:21:12	East bus company	203	BS20001
1000976	31	17 April 2017 11:51:43	West bus company	M409	BS30001

The amount of daily bus GPS data is approximately 9 GB, and there is not much variation in the volume of daily GPS data. The total GPS data volume for a week is approximately 63 GB. The details of GPS data are illustrated in Table 2.

Table 2. GPS data examples.

Sampling Time	Bus Number	Bus Line	Bus Companies	Longitude	Latitude
17 April 2017 18:21:56	BS20008	M443	East bus company	113.825	22.749
17 April 2017 08:09:08	BS20005	M201	East bus company	113.997	22.778
17 April 2017 14:41:43	BS40005	17	West bus company	114.002	22.801
17 April 2017 19:51:53	BS40014	201	West bus company	114.103	22.765

We applied a transit-trip reconstruction approach [44] to reconstruct transit trips based on the aforementioned multi-source data. On weekdays, 3.6 million passenger trips can be recovered per weekday with a volume of 0.67 G. On weekends, 3.1 million trips can be recovered per day with a volume of 0.6 G. The details of the data used can be seen in Table 3, where "#" in transfer field represents the transfer station ("Null" if no transfer happened).

Table 3. Passenger trip examples.

ID	Start Time	Origin Station	Arrival Time	Terminal	Transfer
20001	17 April 2017 10:01:13	Laojie Station	17 April 2017 10:31:23	Luohu Station	Null
20001	17 April 2017 12:21:53	Hongling Road	17 April 2017 13:01:11	Huanggang Station	Null
20005	17 April 2017 17:31:33	Yannan Station	17 April 2017 18:24:47	Meijing Staion	Jingtian Station #
20011	17 April 2017 11:23:44	Laojie Station	17 April 2017 12:44:13	Kanglin Hospital	Museum Station # News Building #

4.2. Transit Mobility Cluster Analysis

The trips at the weekday and weekend were analyzed separately. By using the proposed model, we obtain a compact five-dimensional vector to embed transit mobility for each passenger. Aggregation hierarchy clustering [45] was carried out to extract the transit mobility patterns of passengers. According to the purpose of trips, passengers can be divided into four categories [46]: (1) random: There is only one trip for the passenger; (2) commuting: passengers have two trips a day. The departure time of the first trip should not be later than t_1 (e.g., 10:00 a.m.), while the departure time of the last trip should not be earlier than t_2 (e.g., 6:00 p.m.). It should be noted that t_1 and t_2 should be dynamically adjusted according to different morning and evening peak-hours of different cities. In addition, the origin of the first trip must be same as the destination of the last trip, and the duration between the arrival time of the first trip and the departure time of the last trip must be greater than t_3 (e.g., 8 h); (3) temporary: there are two trips in a day, and the origin of the first trip must be same as the destination of the last trip; and (4) unknown: passengers do not have any of the above three trip patterns. For example, we examined the data of a typical weekday. After trip reconstruction, there are 3,429,852 trips and 1,967,784 passengers. In order to facilitate clustering, we randomly selected 20,000 trips involving 11,851 passengers for cluster analysis. We applied Sum of Squares for Error (SSE) to determine the number of clusters, and set the number of clusters to 5 (Table 2).

As can be seen from the Table 4, cluster 1 contains more commuters making short trips, with a relatively longer average transit duration (exceeds by 5 min), more stops passed, and fewer transfers. Passengers of cluster 1 are characterized as relatively well scheduled, which results in few transfers and long trip times during the morning and evening peak-hours. Cluster 2 has the longest average time, and correspondingly, the highest average number of stops passed and average numbers of transfers. From the perspective of passenger categorization, it has more random passengers. This type of passenger tends to travel long distances and has weak regularity. They are not sensitive to long-distance costs. Cluster 3 and cluster 4 have a lower-than-average time cost and numbers of stops passed, and a slightly high percentage of random passengers than the overall average. Passengers of these two clusters are mainly random passengers who tend to make short-distance trips. Cluster 5 has the fewest passengers with strong randomness and irregular transit patterns. We visualized the transit mobility patterns of the passengers, as depicted in Figure 7. The roads were categorized into three levels based on transit flow volume: high, medium, and low. Additionally, the regions were divided into six levels based on GDP data, with a higher level indicating a higher GDP for that region. In cluster 1, the majority of passengers were found to engage in commuting activities between Luohu and Yantian, as well as between Luohu and Futian. These trips predominantly involved short-distance round trips. During these commutes, frequent transfers were observed at Shiminzhongxin Station and Hongling Station. This commuting behavior often occurred between regions with notable disparities in GDP, such as commercial areas (higher GDP area of Yantian) and residential areas (lower GDP area of Luohu). In cluster 2, approximately 30% of the long-distance passengers showed transit mobility patterns related to Shenzhen North Station. For this group, their transit behavior involved departing from Shenzhen North Station and dispersing to various districts within Shenzhen through typical transfer stations such as Xili Station, Shangmeilin Station, or Buji Station. Due to the prevalence of long-distance journeys, as seen in Figure 7, heavy traffic was observed from Shenzhen North Station until the next transfer station, where transportation pressure was relieved. The visual analysis conducted in this study demonstrates the model's capability to extract deep-lying semantic features related to passenger transit behavior. By analyzing these visual representations, we can gain insights into the dynamics of transit patterns, key transfer stations, and the impact of factors such as GDP disparities on commuting behavior.

Table 4. Passenger transit mobility cluster analysis of a weekday.

Cluster ID	Passenger Count	Ave. Time (min)	Ave. Stops	Ave. Tranfer	Proportion of Passenger Type			
					Random	Commute	Temporary	Unknown
1	2288	40.123	9.877	0.448	0.240	0.307	0.352	0.101
2	1839	42.516	11.183	0.662	0.557	0.214	0.141	0.088
3	5041	33.955	9.137	0.546	0.434	0.196	0.286	0.084
4	2002	34.208	9.266	0.531	0.433	0.202	0.262	0.103
5	681	35.169	9.488	0.543	0.423	0.206	0.268	0.103
Total	11,851	36.587	9.639	0.542	0.415	0.222	0.271	0.092

For comparison, these 11,851 passengers in a weekday were clustered using the original vector representation (as described in Section 3.3.1), as shown in Table 5. There is no significant difference between the clusters. This is because the transit patterns of passengers are too complicated for the original features. While the proposed model targets the characteristics of passengers' public transit mobility, it incorporates more spatio-temporal information, and makes the vector representation more compact. This encourages the representation of passengers with similar transit patterns to be closer together in the latent space. Nevertheless, there is still room for improvement in the absolute accuracy of the results.

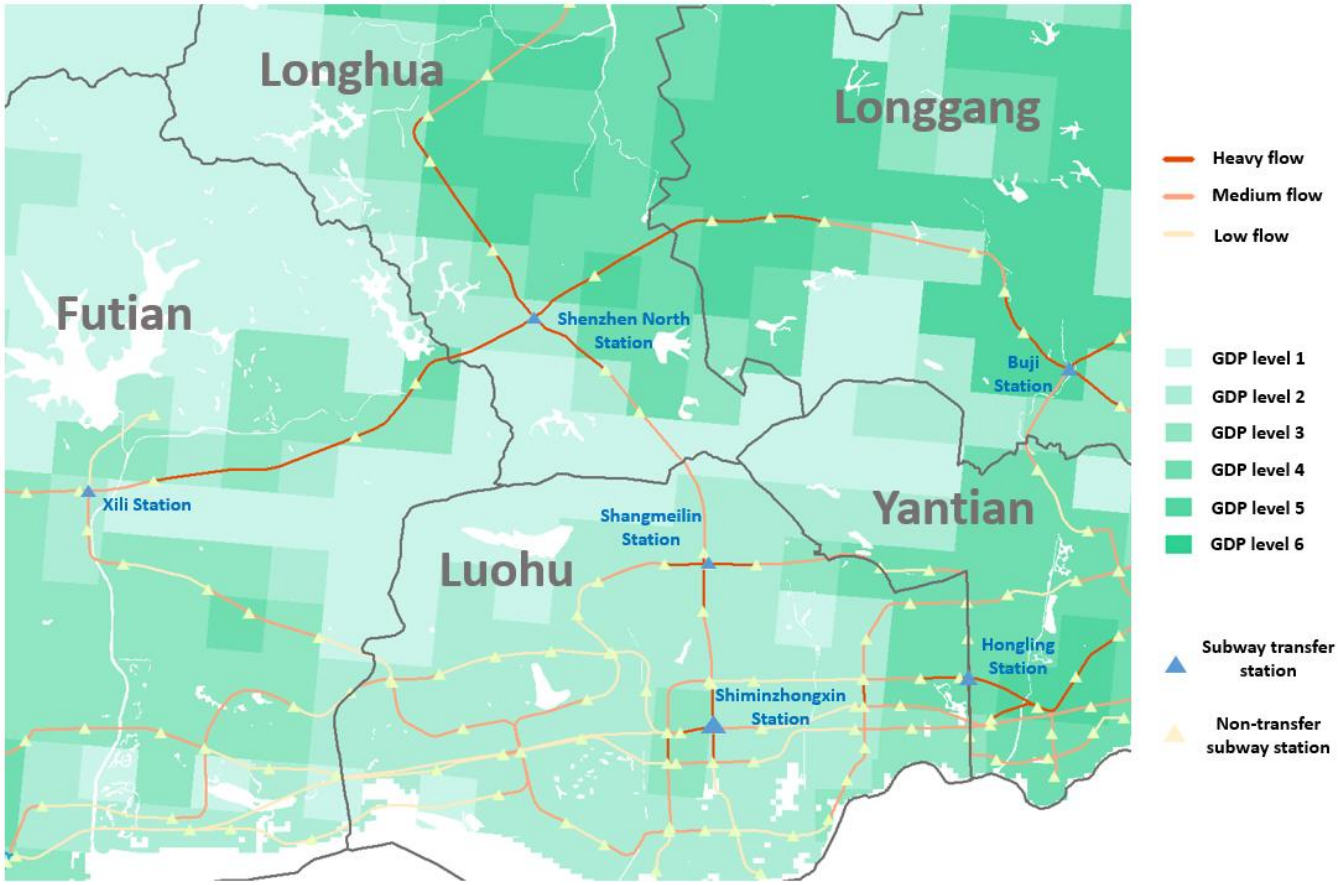

Figure 7. Visualization analysis of a weekday.

Table 5. Original vector cluster analysis of weekday.

Cluster ID	Passenger Count	Ave. Time (min)	Ave. Stops	Ave. Tranfer	Proportion of Passenger Type			
					Random	Commute	Temporary	Unknown
1	6001	36.395	9.568	0.523	0.414	0.233	0.263	0.089
2	1787	36.945	9.722	0.562	0.432	0.206	0.268	0.094
3	2277	36.917	9.773	0.561	0.416	0.196	0.286	0.102
4	768	35.499	9.351	0.541	0.393	0.243	0.274	0.089
5	1018	37.169	9.829	0.573	0.407	0.224	0.288	0.081
Total	11,851	36.587	9.639	0.542	0.415	0.222	0.271	0.092

During the weekend experiment, 3,093,202 trips were used, and 20,000 trips were randomly selected for clustering, consisting of 11,661 passengers (Table 6). The proportion of commuters decreased by about 8% over weekends. Moreover, the proportion of unknown passengers increased by 5%, indicating that weekend passengers have more freedom in their transit behavior, with more trips and higher randomness. There are also passengers with long-distance transit behavior on weekends, as shown in Cluster 1. However, unlike weekdays, such passengers make up a smaller proportion of commuters and short-distance passengers, and a larger proportion of random passengers, indicating that the weekend transit patterns are freer and more complex. In addition, as shown in cluster 2, the distribution of short-distance trip passengers is more concentrated, so this cluster is more compact, indicating that the short-distance transit pattern is more significant on weekends.

Table 6. Passenger transit mobility cluster analysis of weekend.

Cluster ID	Passenger Count	Ave. Time (min)	Ave. Stops	Ave. Tranfer	Proportion of Passenger Type			
					Random	Commute	Temporary	Unknown
1	3506	42.327	11.971	0.632	0.492	0.103	0.172	0.223
2	4233	30.387	8.812	0.491	0.371	0.172	0.349	0.108
3	3015	34.835	9.712	0.552	0.432	0.227	0.212	0.129
4	857	35.208	9.466	0.551	0.411	0.213	0.258	0.118
Total	11,611	35.503	10.047	0.554	0.426	0.168	0.253	0.149

4.3. Estimate Top K Transit Destinations

In this paper, the experiment of estimating the top K transit destinations was also conducted. As shown in Equation (6). Given trips features $x \in X$ and the destination set Q, we solve a subset $Y \subseteq Q$ containing K elements ($K <$ count(Q)) that satisfies the probability that every element y in Y is larger than every element s in S (S is the complement set of Y for Q).

$$P(\forall y \in Y|x) > P(\forall s \in S|x) \tag{6}$$

The following four models were designed for comparison: (1) the proposed model (Ours); (2) the proposed model without the pre-trained sub-generator (OursD) to experiment the impact of the sub-generator; (3) the classical auto-encoder (AE), whose encoder and decoder are composed of GRU and FC, to experiment the impact of the GAN; (4) average aggregation (Raw), which is obtained by average aggregation of personal trip dataset (as described in Section 3.3.1), to compare with the raw features. These four models, as shown in Figure 8, capture different aspects of passenger transit behavior, resulting in differences when used as input features for estimating transit destinations.

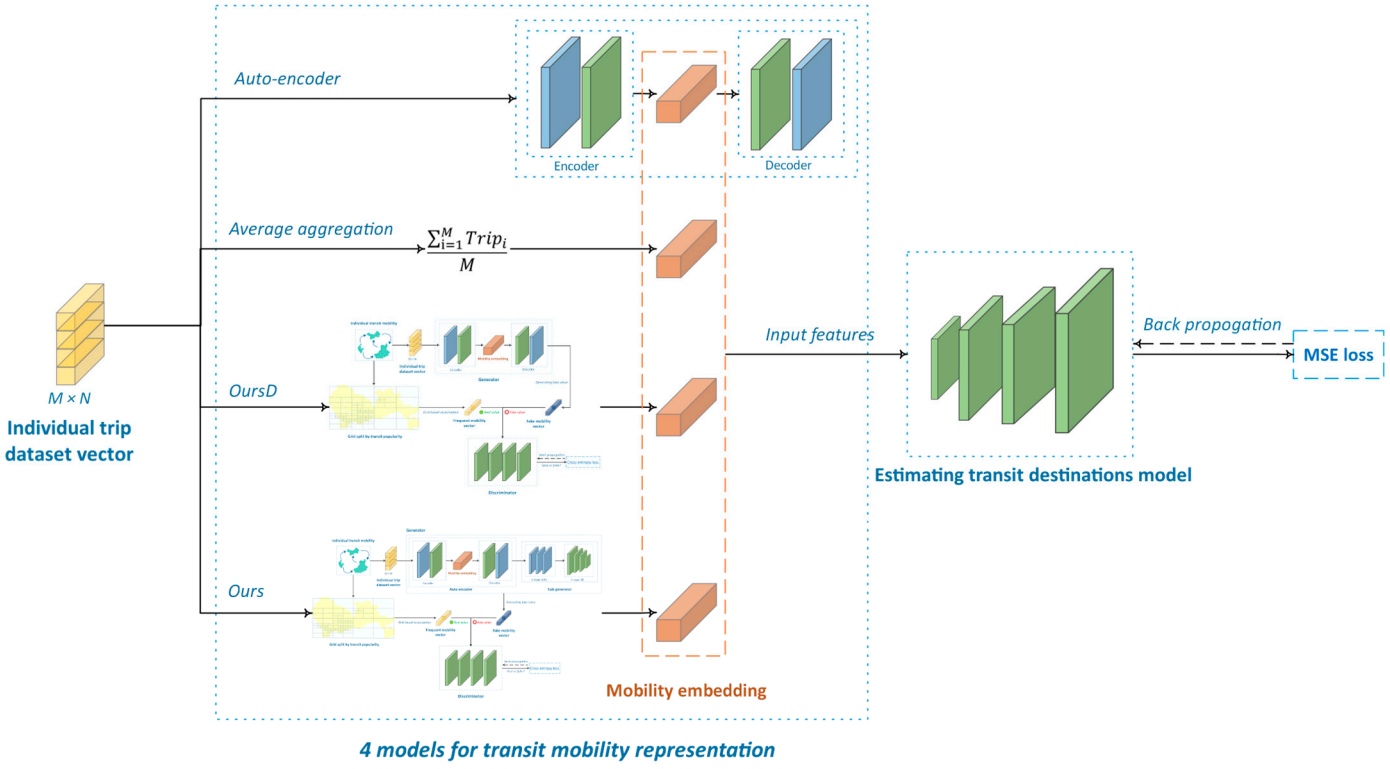

Figure 8. Models for estimating transit destinations.

To estimate the top K transit destinations, we build a model consisting of four FC layers with sigmoid activation functions, as shown in Figure 8. The dimension of each FC layer is 20, 40, and 70, respectively, and the multi-class cross-entropy loss in Equation (7) is

used as the loss function for this model, which x means training data and y means label. K training samples are created for the same passenger to represent the most common K transit destinations. At the same time, in order to reflect the difference in the frequency of the top K transit destinations, the training data set is constructed according to the ratio K:$K-1$:$K-2$:. . .:2:1 during the training process. That is, for the destinations with a higher transit probability, the proportion of the training samples also increases accordingly.

$$CEL(x, y) = -\log \frac{ye^x}{\sum_{j=1}^{n} e^{x_j}} \tag{7}$$

The average relevancy (AR) and average precision (AP), according to the relevant index of the ranking models, are adopted to evaluate the results by Equations (8)–(10), where N is the total number of passengers, K is the count of top transit destinations, y means the real value and y' means the predictive value. $I_y(x)$ function is an indicator function to determine whether a certain element x exists in a set y.

$$AR = \frac{\sum_{i=1}^{N} \frac{\sum_{j=1}^{k} I_y\left(y'_j\right)}{K}}{N} \tag{8}$$

$$AP = \frac{\sum_{i=1}^{N} \frac{\sum_{j=1}^{k} I_{y_j}\left(y'_j\right)}{K}}{N} \tag{9}$$

$$I_y(x) = \begin{cases} 1, & x \in y \\ 0, & x \notin y \end{cases} \tag{10}$$

As shown in Table 7, we evaluated the models with parameter K = 3 and K = 5 for comparison. The results show that there is little difference between Raw and AE. The model without a pre-trained sub-generator (OursD) performs relatively better. The proposed model (Ours) performs the best and has a slight improvement over OursD. Further analysis reveals that the effect of the auto-encoder is mainly in the compression of features. Without spatio-temporal information, the performance of the model is almost the same as using the original features. The proposed model uses a generative adversarial network to fit the real distribution of the data and incorporates spatio-temporal constraints into the model. For this reason, Ours and OursD perform better than Raw and AE.

Table 7. Estimating top K transit destinations.

Model	Count of Transit Destinations			
	K = 3		K = 5	
	AR	**AP**	**AR**	**AP**
Raw	0.626	0.553	0.595	0.527
AE	0.637	0.551	0.609	0.531
OursD	0.707	0.594	0.674	0.562
Ours	**0.723**	**0.637**	**0.709**	**0.581**

5. Conclusions

We have proposed a novel deep learning model that utilizes a generative adversarial network to analyze and understand public passenger transit patterns based on real-world SCD and public transportation network data. Our approach constructs a low-dimensional and compact transit mobility embedding, capturing both frequent transit mobility (representing personal transit characteristics) and similar transit group mobility (representing group transit characteristics). These two characteristics effectively reflect the frequent and regular nature of public transit behavior.

To extract significant mobility patterns for specific public transit passengers, we extended the classic Apriori algorithm by incorporating spatio-temporal constraints. And we

have identified similar transit groups among passengers by considering the constraints of time, space, and semantics. Additionally, given the sparsity of public transit trajectories, we have designed a quad-divided grid strategy based on popularity and area. This approach uses split grids as the basic statistical unit of transit mobility, enabling the conversion of variable-length trips into a unified vector representation and reflecting the flow between different regions.

We utilize a generative adversarial network to aggregate personal transit characteristics and group transit characteristics through an adversarial process, thus embedding a specific public transportation passenger's transit mobility. Through clustering analysis and visual analysis, we confirmed that this embedding captures more transit mobility differences between weekdays and weekends and can detect transit mobility patterns with semantic information, outperforming raw features of individual trips. We also conducted an experiment to estimate the top K transit destinations, which demonstrated the applicability of our proposed method to other human mobility models and yielded favorable results in comparison experiments.

The proposed transit mobility embedding provides a compact representation that enables urban decision-makers and residents to gain insights into the mobility patterns of individuals within a city and the daily operation of the urban transportation system. Notably, we observed a significant number of long-distance travelers moving from Shenzhen North Station to various districts in Shenzhen City. This movement during peak hours in the morning and evening substantially increases the operational pressure on relevant subway lines. Transit decision-makers can use the insights gained from our research to evaluate the current transit systems and develop optimization strategies for the transit network to alleviate the burden on public transportation.

Our research findings also contribute to public transit planning and smart public transportation systems. By leveraging the proposed approach, we can better understand public transit behaviors and inter-urban migration patterns across space and time. Transit planners can gain a deeper understanding of daily transit flow within the city, facilitating the evaluation of system performance and equity. Furthermore, the proposed method aids in the detection of urban functional areas, enabling the improvement of public transit services by adjusting schedules to meet the significant mobility demand in different urban functional areas. However, there is room for improvement in our deep learning model. Future work will focus on model optimization and the incorporation of additional socioeconomic data, such as housing prices and land use, to enhance the modeling and feature engineering processes. These enhancements will contribute to improving the accuracy and completeness of the model's results.

Author Contributions: Conceptualization, Yicong Li and Tong Zhang; methodology, Yicong Li and Tong Zhang; software, Yicong Li, Xiaofei Lv, Yingxi Lu and Wangshu Wang; validation, Yicong Li, Xiaofei Lv, Yingxi Lu and Wangshu Wang; formal analysis, Yicong Li, Xiaofei Lv and Yingxi Lu; investigation, Yicong Li, Xiaofei Lv and Yingxi Lu; resources, Tong Zhang; data curation, Xiaofei Lv, Yingxi Lu and Wangshu Wang; writing—original draft preparation, Yicong Li; writing—review and editing, Tong Zhang; visualization, Xiaofei Lv and Yingxi Lu; supervision, Tong Zhang; project administration, Yicong Li; funding acquisition, Tong Zhang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key R&D Program of China (International Scientific & Technological Cooperation Program), grant number 2019YFE0106500 and the open fund of Wuhan University—Huawei Geoinformatics Innovation Laboratory.

Data Availability Statement: The data presented in this study are unavailable due to privacy or ethical restrictions.

Acknowledgments: The numerical calculations in this paper were conducted on the supercomputing system in the Supercomputing Center of Wuhan University.

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

References

1. Luca, M.; Barlacchi, G.; Lepri, B.; Pappalardo, L. A survey on deep learning for human mobility. *ACM Comput. Surv.* **2021**, *55*, 1–44. [CrossRef]
2. Wang, R.; Li, N.; Wang, Y. Does the returners and explorers dichotomy in urban human mobility depend on the observation duration? An empirical study in Guangzhou, China. *Sustain. Cities Soc.* **2021**, *69*, 102862. [CrossRef]
3. Barbosa, H.; Barthelemy, M.; Ghoshal, G.; James, C.R.; Lenormand, M.; Louail, T.; Menezes, R.; Ramasco, J.J.; Simini, F.; Tomasini, M. Human mobility: Models and applications. *Phys. Rep.* **2018**, *734*, 2018. [CrossRef]
4. Feng, J.; Yang, Z.; Xu, F.; Yu, H.; Wang, M.; Li, Y. Learning to simulate human mobility. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event, CA, USA, 6–10 July 2020; pp. 3426–3433.
5. Sáenz, F.T.; Arcas-Tunez, F.; Muñoz, A. Nation-wide touristic flow prediction with Graph Neural Networks and heterogeneous open data. *Inf. Fusion* **2023**, *91*, 582. [CrossRef]
6. Zipf, G.K. The P_1P_2/D hypothesis: On the intercity movement of persons. *Am. Sociol. Rev.* **1946**, *11*, 677–686. [CrossRef]
7. Simini, F.; González, M.C.; Maritan, A.; Barabási, A.L. A universal model for mobility and migration patterns. *Nature* **2012**, *484*, 96–100. [CrossRef]
8. Masucci, A.P.; Serras, J.; Johansson, A.; Batty, M. Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Phys. Rev. E* **2013**, *88*, 22812. [CrossRef]
9. Xue, H.; Voutharoja, B.P.; Salim, F.D. Leveraging language foundation models for human mobility forecasting. In Proceedings of the 30th International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 1–4 November 2022; Volume 90, pp. 1–9.
10. McKenzie, G.; Romm, D. Measuring urban regional similarity through mobility signatures. *Comput. Environ. Urban Syst.* **2021**, *89*, 101684. [CrossRef]
11. Wu, J.; Qu, Y.; Sun, H.; Yin, H.; Yan, X.; Zhao, J. Data-driven model for passenger route choice in urban metro network. *Phys. A-Stat. Mech. Its Appl.* **2019**, *524*, 787–798. [CrossRef]
12. Ma, X.; Liu, C.; Wen, H.; Wang, Y.; Wu, Y.J. Understanding commuting patterns using transit smart card data. *J. Transp. Geogr.* **2017**, *58*, 135–145. [CrossRef]
13. Rosa, L.; Silva, F.; Analide, C. Explainable Artificial Intelligence on Smart Human Mobility: A Comparative Study Approach. In *Distributed Computing and Artificial Intelligence, Special Sessions, 19th International Conference*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 93–103. [CrossRef]
14. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. In Proceedings of the International Conference on Learning Representations, Scottsdale, AZ, USA, 2–4 May 2013.
15. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Distributed representations of words and phrases and their compositionality. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 3111–3119.
16. Wang, P.; Fu, Y.; Xiong, H.; Li, X. Adversarial Substructured Representation Learning for Mobile User Profiling. In Proceedings of the 25th ACM SIGKDD International Conference, Anchorage, AK, USA, 4–8 August 2019; ACM: New York, NY, USA, 2019.
17. Gao, Q.; Zhou, F.; Zhang, K.; Trajcevski, G.; Luo, X.; Zhang, F. Identifying human mobility via trajectory embeddings. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 1689–1695.
18. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial nets. In Proceedings of the NIPS, Montreal, QC, Canada, 8–13 December 2014.
19. Mauro, G.; Luca, M.; Longa, A.; Lepri, B.; Pappalardo, L. Generating mobility networks with generative adversarial networks. *EPJ Data Sci.* **2022**, *11*, 58. [CrossRef] [PubMed]
20. Agrawal, R.; Imielinski, T.; Swami, A. Mining association rules between sets of items in large databases. *ACM SIGMOD Rec.* **1993**, *22*, 207–216. [CrossRef]
21. Yang, D.; Fankhauser, B.; Rosso, P.; Cudre-Mauroux, P. Location prediction over sparse user mobility traces using RNNs: Flashback in hidden states. In Proceedings of the 29th International Joint Conference on Artificial Intelligence, Yokohama, Japan, 7–15 January 2021; pp. 2184–2190.
22. Barabási, A. The origin of bursts and heavy tails in human dynamics. *Nature* **2005**, *435*, 207–211. [CrossRef]
23. Brockmann, D.; Hufnagel, L.; Geisel, T. The scaling laws of human travel. *Nature* **2006**, *439*, 462–465. [CrossRef] [PubMed]
24. González, M.; Hidalgo, C.; Barabási, A. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779–782. [CrossRef]
25. Song, C.; Qu, Z.; Blumm, N.; Barabási, A.L. Limits of Predictability in Human Mobility. *Science* **2010**, *327*, 1018–1021. [CrossRef]
26. Ellis, D.; Sommerlade, E.; Reid, I. Modelling pedestrian trajectory patterns with Gaussian processes. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), Kyoto, Japan, 27 September–4 October 2009; IEEE: Piscataway, NJ, USA, 2009.
27. Schneider, C.M.; Belik, V.; Couronné, T.; Smoreda, Z.; González, M.C. Unravelling daily human mobility motifs. *J. R. Soc. Interface* **2013**, *10*, 20130246. [CrossRef]
28. Hung, C.C.; Peng, W.C.; Lee, W.C. Clustering and aggregating clues of trajectories for mining trajectory patterns and routes. *VLDB J.* **2015**, *24*, 169–192. [CrossRef]
29. Bagchi, M.; White, P.R. The potential of public transport smart card data. *Transp. Policy* **2005**, *12*, 464–474. [CrossRef]

30. Pelletier, M.P.; Martin, T.; Morency, C. Smart card data use in public transit: A literature review. *Transp. Res. Part C Emerg. Technol.* **2011**, *19*, 557–568. [CrossRef]
31. Ochi, M.; Nakashio, Y.; Yamashita, Y.; Sakata, I.; Asatani, K.; Ruttley, M.; Mori, J. Representation learning for geospatial areas using large-scale mobility data from smart card. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, Heidelberg, Germany, 12–16 September 2016; pp. 1381–1389.
32. Kieu, L.M.; Bhaskar, A.; Chung, E. A modified Density-Based Scanning Algorithm with Noise for spatial travel pattern analysis from Smart Card AFC data. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 193–207. [CrossRef]
33. Yao, D.; Zhang, C.; Zhu, Z.; Huang, J.; Bi, J. Trajectory clustering via deep representation learning. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; IEEE: Piscataway, NJ, USA, 2017.
34. Zhou, F.; Gao, Q.; Trajcevski, G.; Zhang, K.; Zhong, T.; Zhang, F. Trajectory-user linking via variational autoencoder. In Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 13–19 July 2018; pp. 3212–3218.
35. Gao, Q.; Zhou, F.; Trajcevski, G.; Zhang, K.; Zhong, T.; Zhang, F. Predicting Human Mobility via Variational Attention. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 2750–2756.
36. Chen, C.; Liao, C.; Xie, X.; Wang, Y.; Zhao, J. Trip2Vec: A deep embedding approach for clustering and profiling taxi trip purposes. *Pers. Ubiquitous Comput.* **2018**, *23*, 5366. [CrossRef]
37. Cao, H.; Xu, F.; Sankaranarayanan, J.; Li, Y.; Samet, H. Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population. *IEEE Trans. Mob. Comput.* **2019**, *19*, 1096–1108. [CrossRef]
38. Lin, K.; Chen, J.; Lian, X.; Mai, W.; Guo, Z.; Chen, X.; Hsu, T.Y. Spatial-Temporal Context-Aware Location Prediction Based on Bidirectional Self-Attention Network. In Proceedings of the 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, China, 14–17 October 2022; pp. 701–706. [CrossRef]
39. Wang, P.; Fu, Y.; Zhang, J.; Li, X.; Lin, D. Learning Urban Community Structures: A Collective Embedding Perspective with Periodic Spatial-temporal Mobility Graphs. *ACM Trans. Intell. Syst. Technol.* **2018**, *9*, 1–28. [CrossRef]
40. Liu, H.; Li, T.; Hu, R.; Fu, Y.; Gu, J.; Xiong, H. Joint Representation Learning for Multi-Modal Transportation Recommendation. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 1036–1043. [CrossRef]
41. Martin, H.; Wiedemann, N.; Reck, D.J.; Raubal, M. Graph-based mobility profiling. *Comput. Environ. Urban Syst.* **2023**, *100*, 101910. [CrossRef]
42. Gao, Q.; Zhang, F.; Yao, F.; Li, A.; Mei, L.; Zhou, F. Adversarial Mobility Learning for Human Trajectory Classification. *IEEE Access* **2020**, *8*, 20563–20576. [CrossRef]
43. Mizuno, T.; Fujimoto, S.; Ishikawa, A. Generation of individual daily trajectories by GPT-2. *Front. Phys.* **2022**, *10*, 1118. [CrossRef]
44. Zhang, T.; Li, Y.; Yang, H.; Cui, C.; Li, J.; Qiao, Q. Identifying primary public transit corridors using multi-source big transit data. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 1137–1161. [CrossRef]
45. Zepeda-Mendoza, M.L. Resendis-Antonio, O. Hierarchical Agglomerative Clustering. *Encycl. Syst. Biol.* **2013**, *43*, 886–887.
46. Primerano, F.; Taylor, M.A.; Pitaksringkarn, L.; Tisato, P. Defining and understanding trip chaining behaviour. *Transportation* **2008**, *35*, 55–72. [CrossRef]