A Data-Driven Feature Based Learning Application to Detect Freeway Segment Traffic Status Using Mobile Phone Data

Qiang Liu 1,2, Jianguang Xie 1, and Fan Ding 3,*

1 College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; lq@sinoroad.com (Q.L.); xiejg@nuaa.edu.cn (J.X.)
2 Jiangsu Sinoroad Engineering Research Institute Co., Ltd., Nanjing 211008, China
3 School of Transportation, Southeast University, Nanjing 211189, China
* Correspondence: fding5@seu.edu.cn

Abstract: With the finishing of the construction of the main body of a freeway network, adequately monitoring the traffic status of the network has become an urgent need for both travelers and transportation operators. Various methods are proposed to collect traffic information for this purpose. In this article, a data-driven feature-based learning application is implemented to detect segment traffic status using mobile phone data, building on the practical success of deep learning models in other fields. The traffic status estimation is achieved via the application of a three-level long, short-term memory model. Two phone features are extracted from the raw mobile phone data. A large-scale field experiment was conducted using actual data in Jiangsu, China collected over the “National Holiday Golden Week” of 2014. To evaluate the performance, both precision and recall scores are given along with the overall accuracy. The final results of the large-scale experiment indicate that the proposed application performed well and can be an emerging solution for traffic state monitoring when only limited roadside sensing devices are installed.

Keywords: mobile phone data; deep learning model; feature extraction; traffic status detection

1. Introduction

With the finishing of construction on the main body of a freeway network, the ability to fully monitor the traffic status of the network has become an urgent need for both travelers and transportation operators alike [1–3]. Meanwhile, a system that could fully monitor the traffic status would provide the opportunity for further active traffic control and management. Therefore, smart road, which has the ability to communicate with passing vehicles, provides information, and achieves real-time monitoring, has developed rapidly and drawn considerable attention by the public recently [4]. Such need and opportunities are pressing for busy freeways installed with only limited monitoring devices. Regarding the duration and costs for the deployment of a monitoring system, other than conventional solutions (e.g., loop detectors and microwave detectors), new technologies, focusing on emerging big data resources and data mining approaches through machine learning, have been introduced in recent decades [5,6].

Data, generated by wireless communication networks, also known as cellular data, have been used to identify the characteristics of transportation, such as travel demand and traffic details. Regarding the benefits, broad coverage and high penetration rate of the communication network, especially since the internet has gone mobile, have revealed the pattern of individual human mobility from mobile phone data [7,8]. Studies have also indicated that mobile phone location data are a feasible resource for estimating origin-destination flows and travel demand [9–11]. These solutions have proved the potential and opportunities for cellular data application on large-scale issues which do not requiring extreme positioning accuracy.
With the promotion of the big data perspective, a data-driven learning application, which can reveal high-dimensional relations and reduce background noise, is proposed to estimate traffic status using the whole set of mobile phone data, rather than HO data. Two features are extracted from cellular activity data to replace the monitoring parameters generated from HO data, and these features are normalized to partially reduce the background noise. Furthermore, a large-scale experiment was conducted on a busy, 250 km expressway in China. The results were validated by ten bi-directional microwave detector stations. To the best of our knowledge, it was the first experiment with such extensive coverage and that many validation points.

The rest of this paper is organized as follows. In Section 2, a general overview of previous, related studies is provided. Section 3 introduces the composition of the mobile phone data and the procedure of generating two proposed phone features. In Section 4, the learning models applied in this study are presented in detail. The following section, Section 5, presents the experiment and the performance evaluations of this research. Finally, Section 6 concludes that the learning application, based on cellular activity data, performs well in large-scale traffic status estimation.

2. Literature Review

Unlike GPS data, cellular data usually do not contain precise locations, which is a major obstacle for using it to estimate or predict short-term traffic details. In order to determine a phone in a communication network accurately, researchers introduced the handover (HO) mechanism, also known as handoff, which occurs mostly when a phone has an active call and passes the boundary of two cells [12]. Through genuine cellular network HO data, Bar-Gera [13] discussed the possibility of implementing a mobile phone-based system with the function of traffic speeds and travel time estimation. Demissie et al. [14] studied the relation between traffic volumes and the number of HOs over the hours of a day. Evaluations confirmed the positive results and the promising future of HO-based traffic estimation methods [13,15].

However, real communication data always includes a massive number of unexpected and indeterministic errors and interferences, which can affect the performance enormously. Therefore, some researchers tried to fuse HO data with the conventional data of the on-road device, such as loop detectors or microwave detectors, to estimate traffic speed [16], and analyze the performance of the integrated system [17]. On the other hand, other researchers have focused on applying methodologies from different perspectives to reduce the impacts of system noise. Cheng et al. [18] proposed particle filter-based models to cope with system noise and measurement errors. Demissie et al. [14] utilized a single layer backpropagation artificial neural network (BP-ANN) to estimate three discrete levels of arbitrarily defined traffic counts with regard to the noise and sensitivity of HO data. Beside implementing complex statistical and machine learning models, Gundlegård and Karlsson [19] showed that multiple road test runs would help to determine the relatively accurate HO location, thereby mapping with the road network. Nevertheless, multiple tests run for each HO point would increase the cost of deployment for the system of a large-scale freeway network.

Besides the problem of noise, the sample size of HO data is another essential attribute that affects the performance of a phone-based system, much the same as other data sources [20]. Regarding the mechanism of HO data generation, when a phone is on a call and simultaneously passes the boundary of two cells, the HO data is only a small portion of the whole mobile phone dataset. To increase the sample size, Caceres et al. [21] proposed models to estimate traffic flow using in-motion phone data, which is derived by combining HO data and the data of phones having two consecutive calls in two cells. However, such an increase might not be significant, especially considering the fact that many countries forbid the use of mobile phones while driving, phone conversations made while driving typically do not last long enough to pass through two cells, and the number of HO points adjacent to freeways could be sparse for a large-scale freeway. Literature on applying mobile phone data to estimate traffic measures can be found through references [22–25].
3. Cellular Activity Data and Extracted Features

The whole mobile phone dataset includes both network signals, such as location updates (LU) and handover (HO), and records for billing, such as phone call started, phone call ended, text messaging, and data service. Previous studies have usually focused on estimating traffic status using HO data [13,14,18,20]. In the cellular data of this study, raw HO records were only 12.7% of the total dataset. Intuitively, validated and useful HO records are even smaller after filtering and processing. In order to enlarge the sample size, Caceres et al. [21] proposed the combination of HO data and the data of phones with two consecutive calls in two cells together as the model input. However, records of phones with active calls occupy 10.1%, which still cannot take the full advantage of the entire cellular data. Regarding the scale issue [14], that is the relation between traffic and the number of phones, happened at different scales for different locations and for large-scale freeway traffic monitoring, the intuitive idea is to involve as many validated phones as possible into estimating traffic status. Therefore, instead of only involving phones with consecutive calls in two cells, phones that had consecutive records within a relatively short duration in two cells which are close to the target freeway were also counted. In this way, the potential data set was enlarged to almost cover the entire raw cellular activity data. Subsequently, Ding et al. [25] established a method to extract two phone features, i.e., phone speed (PSP) and phone count (PC).

In this method, the base station, which is the center of each cell in the communication network, is mapped to the freeway based on the projection point and assigned with accumulative mileage computed from the route origin, which allows for a unique mileage for an individual base station. Regarding each raw data record only containing the base station identification, the projection point of the corresponding base station of each record is used to position the phone and is represented as $m_{\text{record}}$. The flowchart of the phone speed feature extraction for a certain phone, $p$, in an interval of $t$ of a certain segment is shown in Figure 1.

As shown in Figure 1, $r_{p,t}$ and $r_{p,t-1}$ represent all raw mobile phone data records of a certain phone, $p$, during time interval $t$ and $t-1$ in the target segment, respectively.

The phone $p$ was defined as a valid phone if the mileage difference between the first and last records, in chronological order, during interval $t$ in $r_{p,t}$, $m_{\text{first}}(r_{p,t})$ and $m_{\text{last}}(r_{p,t})$ of the phone, $\Delta m_{p,t}$, was not equal to zero. In addition, the movement direction of the phone could be determined through the sign of $\Delta m_{p,t}$. The phone speed of phone $p$ was acquired via the computation of the ratio of the absolute value of $\Delta m_{p,t}$ to the time difference between the first and last records of phone $p$ during interval $t$, $\Delta t_{p,t}$.

On the other hand, if $\Delta m_{p,t}$ was equal to zero, the method retrieved the first record of phone $p$ in its previous interval, i.e., $m_{\text{first}}(r_{p,t-1})$, and then used the abovementioned algorithm to calculate the phone speed again.

However, if $\Delta m_{p,t}$ matches none of the two cases mentioned above, the phone $p$ was defined as invalid and filtered out.

As indicated by Ding et al. [25], the moving direction availability is the filter condition for the phones. The phone count can be kept as a second feature if the phone’s moving direction is available for each direction. Finally, the phone speed of each segment $i$ over interval $t$ in a certain direction, $ps_{i,t}$, can be generated by calculating the mean value of the speed features of all valid phones in the corresponding direction and interval, and the phone count of each segment $i$ over interval $t$, $pc_{i,t}$, can be generated through counting the number of valid phones for each direction.
Due to the relaxed filter restriction, the scale issue, and the system noise, features were also applied with a scaling normalization method to regularize the searching space. The model described in the next section was applied not only to reveal the pattern of estimating traffic status, but also to further denoise.

4. Learning Models

Based on those two extracted features, learning models were implemented to estimate the traffic status. In recent years, neural networks showed outstanding performance in practical contests [26]. Regarding the spatiotemporal nature of traffic status among segments, a unique recurrent neural network (RNN) architecture, the long short-term memory (LSTM), was implemented since it has the capability of revealing the optimal time lags automatically. The detailed structure of an LSTM block is shown in Figure 2.
The LSTM was proposed by Hochreiter and Schmidhuber [27], and it was designed to avoid the gradient vanishing issue [28]. An LSTM block consists of a memory cell, \( c \), and three gates: the input gate, \( i_t \), the output gate, \( o_t \), and the forget gate, \( f_t \). The basic computation workflow of a block for interval \( t \) can be expressed in the form as follows:

\[
\begin{align*}
    i_t &= \text{sigmoid}(W_{in,x}x_t + W_{in,m}m_{t-1} + W_{in,c}c_{t-1} + b_{in}) \\
    f_t &= \text{sigmoid}(W_{f,x}x_t + W_{f,m}m_{t-1} + W_{f,c}c_{t-1} + b_f) \\
    z_t &= \tanh(W_{z,x}x_t + W_{z,m}m_{t-1} + b_z) \\
    c_t &= f_t \odot c_{t-1} + i_t \odot z_t \\
    o_t &= \text{sigmoid}(W_{o,x}x_t + W_{o,m}m_{t-1} + W_{o,c}c_t + b_o) \\
    m_t &= o_t \odot \tanh(c_t)
\end{align*}
\]

Here, \( \odot \) denotes element-wise multiplication (Hadamard product), \( x_t \) denotes the input of the block over interval \( t \), \( W \) denotes the weight matrices between corresponding vectors (e.g., \( W_{in,x} \) represents the matrix of weights from the input gate, \( x_t \), to the input, \( i_t \), of a memory cell), \( b \) represents the vector of bias values (e.g., \( b_f \) represents the bias at the forget gate, \( f_t \), of layer \( k \)), and \( m_t \) denotes the output of the LSTM memory block over interval \( t \). The activation function of all gates is the sigmoid function [29], and the tanh-function is applied for other activations.

With the trend of the model going “deeper” in the neural network field, studies have confirmed that a deep learning model could give an excellent performance on pattern recognition, classification, and prediction-related topics, though theoretical interpretability is an on-going task. Regarding the high background noise of the raw mobile phone data, and the advantage of figuring out the high dimensional relations among features, a three-level LSTM neural network (LSTM-3) was implemented for this study (Figure 3).

**Figure 3.** The structure of the proposed LSTM-3 network.

For each direction of segment \( i \) over interval \( t \), the input of layer \( k \) is denoted as \( x_t^{k,i} \), while the output is denoted as \( y_t^{k,i} \). Specifically, the definitions are as follows:

\[
\begin{align*}
    x_t^{k,i} &= \begin{cases} 
        (ps_t^i, pc_t^i)^T & k = 1 \\
        \text{otherwise} & \text{otherwise}
    \end{cases} \\
    y_t^{k,i} &= l\left(W_{y,m}m_t^{k,i} + b_y^k\right)
\end{align*}
\]

Here, \( ps_t^i \) and \( pc_t^i \) are the corresponding feature of segment \( i \) over interval \( t \). \( W_{y,m} \), \( m_t^{k,i} \), and \( b_y^k \) are the weight, the output vector of LSTM blocks, and the bias, respectively. A linear function, \( l(\cdot) \), is used to aggregate attributes.
The final output was the traffic status. The traffic status here classified traffic conditions into three levels: free flow (L0), transition mode (L1), and severe congestion (L2). Each level was assigned a unique color (Table 1). In this study, unlike the work of Demissie et al. [14], the levels were uniformly divided based on traffic speed limits rather than on volume. This was because the nature of traffic flow, where the low traffic could either refer to free flow status or severe congestion, cannot represent the running conditions from the operational point of view. The speed range of each level was defined arbitrarily by the local traffic operation center.

| Level | Description          | Color   | Speed Range |
|-------|----------------------|---------|-------------|
| L0    | Free flow            | Green   | ≥80         |
| L1    | Transition mode      | Yellow  | (40, 80)    |
| L2    | Severe congestion    | Red     | ≤40         |

Furthermore, a single layer feedforward neural network (FNN-1) and a single layer LSTM neural network (LSTM-1) were also applied to compare the classification performance on the time series data of phone features with the LSTM-3 model.

5. Case Study in Jiangsu, China

The testbed of this study is a busy, eight-lane freeway in Jiangsu, China [30]. It is divided evenly into 250 one-kilometer-long segments. Figure 4 demonstrates the view of the research area. The data collection, for both cellular activity data and microwave data, was conducted from 00:00 1 October 2014 to 23:59 7 October 2014, covering the whole “National Day Golden Week”. The microwave data were collected from roadside detector stations and aggregated by a local transportation operation center for every 5 min. As the detectors of station No. 4 were not fully functional, only data from the other ten stations were involved in this experiment for both directions. To reduce the impacts of outliers and noise, microwave data were firstly smoothed by the seasonal trend decomposition (STL) procedure [31], then classified into three levels based on the speed (Table 1).

Figure 4. General view of the research area and locations of roadside detectors. (Note: background photo is from Google Maps).

Based on the processed microwave data for the ten validation segments of the westbound direction, over the “National Day Golden Week” 19,366 intervals were labeled as L0, 203 intervals as L1, and the remaining 591 intervals as L2. For the eastbound direction, the numbers were 19,146 L0, 429 L1, and 585 L2. The three levels of traffic status were exceptionally unevenly distributed, which fit with expectations. However, this would impact the performance evaluation if only the overall classification accuracy was considered as other studies did [14], since that estimated all intervals at L0 could already achieve high accuracy.
Therefore, two terms, precision and recall, were introduced to evaluate the performance of learners for each traffic state level.

In similar topics, such as pattern recognition and binary classification, precision and recall are widely applied to measure the relevance, particularly for imbalanced classification problems. Precision is the ratio of correctly estimated instances among the retrieved instances, while recall is the fraction of correctly estimated instances over the total amount of actual observed, relevant instances. For a single class, Equations (9) and (10) use true positive (TP, actual true and estimated as true), false positive (FP, actual true and estimated as false), and false negative (FN, actual false and estimated as false) to present the definition of precision and recall.

\[
\text{precision} = \frac{TP}{TP + FP} \quad (9)
\]

\[
\text{recall} = \frac{TP}{TP + FN} \quad (10)
\]

Tables 2 and 3 indicate the confusion matrixes, corresponding precision, recall and overall accuracy for both directions and three learners. After training, all learners could detect three traffic status levels and especially perform well in the detection of free flow status and severe congestions. Compared to the other two levels, all learners did not express promising performance on the transition mode (L1). The potential reasons for this specific phenomenon include:

1. The transition mode is a relatively unstable state and typically only lasts for a short period of time. Therefore, there are not adequate samples to train and validate learning models, especially deep ones;
2. Current validation data were collected from microwave detectors, which report the instantaneous speed for vehicles. For some intervals, the change of levels defined by such speed could be short and intense (jumping L1) even though the aggregation and the denoise processes were applied. However, extracted features were generated by the average space-mean speed of each valid phone, so learners might incorrectly classify the traffic state due to the sensitivity;
3. Though all detectors were fully functional during the research period, no individual microwave detector is identical to another, which means results could vary for different detectors even if monitoring the same traffic condition. It is better to involve a second validation data source to verify the performance and correctness further.

**Table 2.** Confusion matrix and performance of learners for the westbound direction.

| Observed Level | FNN-1 Predicted | LSTM-1 Predicted | LSTM-3 Predicted |
|----------------|-----------------|-----------------|-----------------|
|                | L0   | L1   | L2   | L0   | L1   | L2   | L0   | L1   | L2   |
| L0             | 19,027 | 275 | 64 | 19,066 | 224 | 76 | 19,147 | 159 | 60 |
| L1             | 79    | 102  | 22 | 84    | 102 | 17 | 94    | 77  | 32 |
| L2             | 59    | 244  | 288 | 50    | 283 | 258 | 104   | 220 | 267 |

**Performance**

|                   | FNN-1 | LSTM-1 | LSTM-3 |
|-------------------|-------|--------|--------|
| Precision         | 0.993 | 0.770  | 0.985  |
| Recall            | 0.982 | 0.487  | 0.989  |
| Overall Accuracy  | 0.963 | 0.964  | 0.967  |
Table 3. Confusion matrix and performance of learners for the eastbound direction.

| Observed Level | FNN-1 Predicted | LSTM-1 Predicted | LSTM-3 Predicted |
|----------------|-----------------|-----------------|-----------------|
|                | L0   | L1   | L2   | L0   | L1   | L2   | L0   | L1   | L2   |
| L0             | 18,653 | 464  | 29   | 18,813 | 294  | 39   | 18,860 | 241  | 45   |
| L1             | 83    | 284  | 62   | 132   | 236  | 61   | 106   | 250  | 73   |
| L2             | 19    | 164  | 402  | 24    | 165  | 396  | 38    | 114  | 433  |

Performance

|               | Precision | Recall  | Overall Accuracy |
|---------------|-----------|---------|------------------|
| LSTM-3 Predicted | 0.992     | 0.985   | 0.969            |
| LSTM-1 Predicted | 0.992     | 0.985   | 0.969            |
| FNN-1 Predicted | 0.995     | 0.974   | 0.969            |

In general, learning models could distinguish traffic states based on extracted features from cellular activity data. Furthermore, based on the overall accuracy, the three-level LSTM model showed the best performance for both directions although the differences were not very significant compared to the ones of the other two learners. Nevertheless, this improvement was hard to achieve due to the complex high-dimensional feature relations. The deep model indeed learned to reveal the relationship between traffic states and extracted features. Figure 5 demonstrates the comparisons between levels according to microwave data (after STL procedure) and the LSTM-3 model outputs for validation links with L1 and L2 conditions. As the study covered an entire week, and each interval was 5 min, pie charts were applied to present the results of selected validation links. The left half circle of each chart indicates the traffic status levels based on traffic speed obtained from microwave detectors and the right represents the model outputs.

Figure 5. Comparisons between traffic status levels according to microwave data and LSTM-3 outputs for selected validation links.

Intuitively, the pie charts in Figure 5 are mostly symmetrical. Therefore, the LSTM-3 outputs are essentially matched with the traffic status as defined by the microwave data. There are certain false alarms and detection time latency, which are because of the nature of cellular activity data and the reasons mentioned earlier. Meanwhile, Figure 6 presents the final results of LSTM-3 covering the entire 250 km large-scale freeway over the research period. The diagram matches the practical understanding of bottlenecks on a freeway and the concept of spatiotemporal congestion propagation. According to the final results, the learning model is a promising method to detect three level traffic states under extreme volumes when only limited roadside equipment is installed.
Figure 6. The final results of LSTM-3 for all segments over the research period (a) Westbound results of the testbed freeway (b) Eastbound results of the testbed freeway.

6. Summary and Future Works

This research presents a data-driven learning-based method to detect three level traffic states using a big data resource, specifically, cellular activity data. Two features, the phone count (PC) and the phone speed (PS), were extracted from the raw cellular activity data and considered as model inputs for learners. Three types of learners, a single layer feedforward neural network (FNN-1), a single layer LSTM network (LSTM-1), and a three-level LSTM (LSTM-3), were employed to indicate and verify the performance of learning models, especially the deep model, on traffic state estimation. The proposed method took advantage of the broad coverage of the wireless communication network and used features, rather than conventional handover and location update records, to monitor traffic states.

The case study of this research took place on a busy, 250 km expressway in China during the “National Day Golden Week” of 2014. The three traffic status levels were defined by the local traffic operation center. On the expressway, there were ten working microwave detector stations, which were selected as validation segments, and the ground truth levels were computed through the data collected by the microwave detectors. Regarding the fact that traffic states are not uniformly distributed, two evaluation parameters, precision and recall, as well as overall accuracy, were involved in measuring the estimation results. According to the overall accuracy, the deep model achieved the best outcomes for both directions. Meanwhile, a detailed spatiotemporal traffic states diagram was demonstrated to perform the final results, which indicates the promising performance and the bright future of the proposed method to monitor traffic details for freeways with limited conventional roadside detectors installed.

Nevertheless, the results were promising, but still have space for future improvements. Additional studies are desperately needed, including examining the impacts of traffic states levels, specifically the different numbers of levels and different ranges for each level. Future research could work on proposing methods to detect the transition mode precisely.
and in detail. With further enhancements, the traffic monitoring solution based on cellular activity data should be a promising implementation that could cover large-scale freeways and be deployed easily, even with limited numbers of roadside equipment.

**Author Contributions:** Conceptualization, F.D.; Formal analysis, Q.L.; Methodology, Q.L., J.X. and F.D.; Resources, F.D.; Writing—original draft, Q.L.; Writing—review & editing, J.X. and F.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by the Fundamental Research Funds for the Central Universities (Grant No. 2242021R10039).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available on request due to restrictions e.g., privacy or ethical.

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

**References**

1. Bucknell, C.; Herrera, J.C. A trade-off analysis between penetration rate and sampling frequency of mobile sensors in traffic state estimation. *Transp. Res. Part C Emerg. Technol.* 2014, 46, 132–150. [CrossRef]

2. Herrera, J.C.; Work, D.B.; Herring, R.; Ban, X.; Jacobson, Q.; Bayen, A.M. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transp. Res. Part C Emerg. Technol.* 2010, 18, 568–583. [CrossRef]

3. Wang, W.; Jin, J.; Ran, B.; Guo, X. Large-scale freeway network traffic monitoring: A map-matching algorithm based on low-logging frequency GPS probe data. *J. Intell. Transp. Syst. Technol. Plan. Oper.* 2011, 15, 63–74. [CrossRef]

4. Trubia, S.; Severino, A.; Curto, S.; Arena, F.; Pau, G. Smart Roads: An Overview of What Future Mobility Will Look Like. *Infrastructures* 2020, 5, 107. [CrossRef]

5. Bachmann, C.; Roorda, M.J.; Abdulhai, B.; Moshiri, B. Fusing a bluetooth traffic monitoring system with loop detector data for improved freeway traffic speed estimation. *J. Intell. Transp. Syst. Plan. Oper.* 2013, 17, 152–164. [CrossRef]

6. Tanikella, H.; Smith, B.L. An investigation of the application of stratified sampling in probe-based traffic-monitoring systems. *J. Intell. Transp. Syst. Technol. Plan. Oper.* 2010, 14, 83–94. [CrossRef]

7. Calabrese, F.; Diao, M.; Di Lorenzo, G.; Ferreira, J.; Ratti, C. Understanding individual human mobility patterns from urban sensing data: A mobile phone trace example. *Transp. Res. Part C Emerg. Technol.* 2013, 26, 301–313. [CrossRef]

8. González, M.C.; Hidalgo, C.A.; Barabási, A.-L. Understanding individual human mobility patterns. *Nature* 2008, 453, 779–782. [CrossRef]

9. Caceres, N.; Wideberg, J.P.; Benitez, F.G. Deriving origin–destination data from a mobile phone network. *IET Intell. Transp. Syst.* 2007, 1, 15. [CrossRef]

10. Calabrese, F.; Di Lorenzo, G.; Liu, L.; Ratti, C. Estimating Origin-Destination Flows Using Mobile Phone Location Data. *IEEE Pervasive Comput.* 2011, 10, 36–44. [CrossRef]

11. Wang, M.-H.; Schrock, S.D.; Vander Brock, N.; Mulinazzi, T. Estimating Dynamic Origin-Destination Data and Travel Demand Using Cell Phone Network Data. *Int. J. Intell. Transp. Syst. Res.* 2013, 11, 76–86. [CrossRef]

12. Yilin, Z. Mobile phone location determination and its impact on intelligent transportation systems. *IEEE Trans. Intell. Transp. Syst.* 2000, 1, 55–64. [CrossRef]

13. Bar-Gera, H. Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: A case study from Israel. *Transp. Res. Part C Emerg. Technol.* 2007, 15, 380–391. [CrossRef]

14. Demissie, M.G.; de Almeida Correia, G.H.; Bento, C. Intelligent road traffic status detection system through cellular networks handover information: An exploratory study. *Transp. Res. Part C Emerg. Technol.* 2013, 32, 76–88. [CrossRef]

15. Liu, H.X.; Danczyk, A.; Brewer, R.; Starr, R. Evaluation of Cell Phone Traffic Data in Minnesota. *Transp. Res. Rec.* 2008, 1–7. [CrossRef]

16. He, S.; Zhang, J.; Cheng, Y.; Wan, X.; Ran, B. Freeway Multisensor Data Fusion Approach Integrating Data from Cellphone Probes and Fixed Sensors. *J. Sens.* 2016, 2016, 1–13. [CrossRef]

17. Zhang, J.; He, S.; Wang, W.; Zhan, F. Accuracy Analysis of Freeway Traffic Speed Estimation Based on the Integration of Cellular Probe System and Loop Detectors. *J. Intell. Transp. Syst. Plan. Oper.* 2015, 19, 411–426. [CrossRef]

18. Peng, C.; Zhijun, Q.; Bin, R. Particle filter based traffic state estimation using cell phone network data. In Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference, Toronto, ON, Canada, 17–20 September 2006; pp. 1047–1052. [CrossRef]

19. Gundlegard, D.; Karlsson, J.M. Handover location accuracy for travel time estimation in GSM and UMTS. *IET Intell. Transp. Syst.* 2009, 3, 87. [CrossRef]

20. He, S.; Cheng, Y.; Zhong, G.; Ran, B. A data-driven study on the sample size of cellular handoff probe system. *Adv. Mech. Eng.* 2017, 9, 168781401769844. [CrossRef]
21. Caceres, N.; Romero, L.M.; Benitez, F.G.; del Castillo, J.M. Traffic Flow Estimation Models Using Cellular Phone Data. *IEEE Trans. Intell. Transp. Syst.* 2012, 13, 1430–1441. [CrossRef]

22. Zhang, J.; Wang, F.-Y.; Wang, K.; Lin, W.-H.; Xu, X.; Chen, C. Data-Driven Intelligent Transportation Systems: A Survey. *IEEE Trans. Intell. Transp. Syst.* 2011, 12, 1624–1639. [CrossRef]

23. Caceres, N.; Wideberg, J.; Benitez, F.G. Review of traffic data estimations extracted from cellular networks. *IET Intell. Transp. Syst.* 2008, 2, 179. [CrossRef]

24. Qiu, Z.; Cheng, P. State of the art and practice: Cellular probe technology applied in advanced traveler information system. In Proceedings of the Transportation Research Board 86th Annual Meeting, Washington, DC, USA, 21–25 January 2007.

25. Ding, F.; Zhang, Z.; Zhou, Y.; Xiaoxuan, C.; Bin, R. Large-Scale Full-Coverage Traffic Speed Estimation under Extreme Traffic Conditions Using a Big Data and Deep Learning Approach: Case Study in China. *Transp. Eng. Part A Syst.* 2019, 5, 145. [CrossRef]

26. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* 2015, 61, 85–117. [CrossRef]

27. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* 1997, 9, 1735–1780. [CrossRef]

28. Hochreiter, S.; Bengio, Y.; Frasconi, P.; Schmidhuber, J. Gradient Flow in Recurrent Nets: The Difficulty of Learning Long Term Dependencies. In *A Field Guide to Dynamical Recurrent Networks*; Wiley-IEEE Press: Hoboken, NJ, USA, 2001. [CrossRef]

29. Han, J.; Moraga, C. *The Influence of the Sigmoid Function Parameters on the Speed of Backpropagation Learning*; Springer: Berlin, Germany, 1995; pp. 195–201. [CrossRef]

30. G42 Shanghai–Chengdu Expressway—Wikipedia [WWW Document], n.d. Available online: https://en.wikipedia.org/wiki/G4_2_Shanghai--Chengdu_Expressway (accessed on 31 December 2017).

31. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *J. Off. Stat.* 1990, 6, 3–33.