entropy

MDPI

Article

An Entropy-Based Approach for Evaluating Travel Time Predictability Based on Vehicle Trajectory Data

Tao Xu [1,2], Xianrui Xu [1,2], Yujie Hu [3] and Xiang Li [1,2,*]

1 Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China; txucn@hotmail.com (T.X.); xuxianrui105@163.com (X.X.)
2 School of Geographic Sciences, East China Normal University, Shanghai 200241, China
3 Kinder Institute for Urban Research, Rice University, Houston, TX 77005, USA; yhu@rice.edu
* Correspondence: xli@geo.ecnu.edu.cn; Tel.: +86-21-5434-1218

Academic Editor: Kevin H. Knuth
Received: 29 January 2017; Accepted: 7 April 2017; Published: 11 April 2017

Abstract: With the great development of intelligent transportation systems (ITS), travel time prediction has attracted the interest of many researchers, and a large number of prediction methods have been developed. However, as an unavoidable topic, the predictability of travel time series is the basic premise for travel time prediction, which has received less attention than the methodology. Based on the analysis of the complexity of the travel time series, this paper defines travel time predictability to express the probability of correct travel time prediction, and proposes an entropy-based method to measure the upper bound of travel time predictability. Multiscale entropy is employed to quantify the complexity of the travel time series, and the relationships between entropy and the upper bound of travel time predictability are presented. Empirical studies are made with vehicle trajectory data in an express road section to shape the features of travel time predictability. The effectiveness of time scales, tolerance, and series length to entropy and travel time predictability are analyzed, and some valuable suggestions about the accuracy of travel time predictability are discussed. Finally, comparisons between travel time predictability and actual prediction results from two prediction models, *ARIMA* and *BPNN*, are made. Experimental results demonstrate the validity and reliability of the proposed travel time predictability.

Keywords: travel time predictability; multiscale entropy; travel time series; vehicle trajectory data

1. Introduction

In transportation, travel time is confined to the time traveled by vehicles in a road network. Accordingly, travel time forecast is to predict the time taken by a vehicle to travel between any two points in a road network, which may benefit transportation planning, design, operations, and evaluation [1]. With ever-increasing traffic congestion in metropolitan areas, travel time for every traveler becomes complicated and irregular. How to accurately predict travel time, therefore, is of great importance to researchers. In the past few decades, a variety of travel time prediction approaches has been developed such as the linear regression, *ARIMA*, Bayesian nets, neural networks, decision trees, support vector regression and Kalman filtering methods. Based on periodic fluctuations of historical travel time series, these approaches make use of their own derivation rules to recognize traffic patterns and predict future travel time on specific routes as precisely as possible. However, the inevitable nonstationarity of travel time series caused by high self-adapting and heterogenetic drivers or by unpredictable and unusual circumstances makes it difficult to accurately predict future travel time. In addition to the performance of prediction models, the quality of input data (historical travel time series) for prediction model affects the precision of travel time prediction.

Nowadays, various data are employed for travel time prediction [2–4], e.g., vehicle trajectory data, mobile phone data, smart card data, loop detector data, video monitoring data, and artificial statistical data. Vehicle trajectory data are frequently collected from a large number of vehicles and consist of a huge number of GPS sample points including geographic coordinates and sample time as well as the identification of vehicle. Historical travel time series in specific trips can be extracted from large amounts of vehicle trajectory data, and travel time prediction can be achieved. Based on vehicle trajectory data, many existing research works have been performed for travel time prediction. Some of them evaluate the performance of prediction models [3,5–7], and some of them focus on the reliability or uncertainty of historical travel time series [8–13], but few of them examine the quality of data from the perspective of prediction.

Travel time reliability is the consistency or dependability in travel times as measured day-to-day or at different times of a day [14], which represents the temporal uncertainty experienced by travelers in their trip [15] or the travel time distributions under various external conditions [16]. Travel time reliability only presents the certainty of historical travel rules but not the accuracy of future travel times.

The aim of this paper is to evaluate the effectiveness of historical travel time series extracted from vehicle trajectory data in travel time prediction. Especially, we use the term "predictability" to denote the results of evaluation. In traffic studies, some research efforts about predictability have been presented. Yue et al. [17] used the cross-correlation coefficient between traffic flows collected at two detector stations to explain short-term traffic predictability in the form of probability. Foell et al. [18] analyzed the temporal distribution of ridership demand on various date conditions and used the F-score, an effective metric of information retrieval, to measure the predictability of bus line usage. Siddle [19] introduced the travel time predictability of two specific prediction models—auto-regressive moving average and non-linear time series analysis—in the Auckland strategic motorway network; and travel time predictability was used to explain the performance of specific prediction models. In addition, the predictability of road section congestion (speed) [20] and human mobility [21,22] are measured by information entropy. Until now, there is not enough literature on data-driven measurement of travel time predictability. Therefore, we hope to explore travel time predictability for evaluating the characteristic of travel time series on prediction.

In this paper, travel time predictability describes the possibility of correct time prediction based on historical travel time series. It indicates the influence of the complexity of historical travel time series on prediction results. For example, travel time predictability of a travel time series being 0.9 indicates that the travel time prediction accuracy cannot exceed 90% for the given travel time series no matter how good a predictive model is. Regular commute patterns give us confidence about future travel times, but random traffic flow often disturbs traffic rules and bring uncertain changes to travel time prediction.

Song et al. [22] explored the limits of predictability of human mobility and developed a method to measure the upper bound of predictability based on information entropy [23] and Fano's inequality [24]. In their research, mobile phone data were employed to quantify human mobility as discrete location series, and the entropy of location series was measured by Lempel-Ziv data compression [25]; Fano's inequality was used to deduce the relationships between entropy and the upper bound of predictability. A Lempel-Ziv data compression algorithm is a method to measure the complexity of the nonlinear symbolic coarse-grained time series. However, travel time series usually has a continuous range of values determined by the tradeoff between accuracy and grain size. Since the complexity of time series is highly sensitive to grain size [26], the symbolization of travel time series is never a straightforward task. Furthermore, the complexity of the Lempel-Ziv algorithm can only be used for qualitative analysis and is not suitable for quantitative description [27]. Therefore, the Lempel-Ziv algorithm is not suitable for measuring the complexity of travel time series, and the method proposed by Song et al. [22] is not applicable to travel time predictability.

Inspired by Song et al. [22], this paper attempts to measure the complexity of travel time series and assess travel time predictability. First, a travel time series is defined as a continuous variable,

and multiscale entropy (*MSE*) [28] in different scales is measured to present the true entropy of the given travel time series. Then, the upper bound of predictability is calculated based on the method of Song et al. [22]. Usually, *MSE* is used to assess the complexity of multi-value time series from the perspective of multi-time scales and has been successful in many fields. However, *MSE* often produces some inaccurate estimations or undefined entropy which could largely bias the evaluation of the complexity of travel time series. Wu et al. [29] proposed the refined composite multiscale entropy (*RCMSE*) algorithm based on *MSE*, and found that *RCMSE* increases the accuracy of entropy estimation and reduces the generation of undefined entropy.

To this end, our contributions are to integrate the two methods proposed by Wu et al. [29] and Song et al. [22], and apply them to evaluate the features of entropy and predictability of travel time series.

This paper applies the above techniques to an express road section with heavy traffic flow in Shanghai, China. Taxi cab trajectory data recorded in April 2015 are used to present the average taxi mobility trends and evaluate travel time predictability. Commonly seen mobility research often focuses on the mobility patterns of independent individuals (person or vehicle) based on such as mobile phone data [22,30,31] and GPS data [22,32,33]. Differently, travel time predictability emphasizes the expected success rate of travel time prediction based on a given travel time series from individual or statistics values of travel time, and the statistical trends of travel times from multiple vehicles may present traffic patterns more effectively than individual mobility which may be affected by unpredictable driving behaviors. In this paper, massive trip data in one route are acquired from a large amount of taxi trajectory data, and the 5 min travel time series is averaged to present the statistical trends of travel time. We then assess the entropy and predictability based on the travel time series. Next, we discuss the influences of time scales, tolerance, and series length on entropy and travel time predictability. Finally, we employ two prediction models, *ARIMA* and *BPNN*, to predict the future travel time of the selected route for model validation.

The rest of this paper is organized as follows. The next section defines the methodology of travel time predictability. The study area, data sources and results of a case study are presented in Section 3. Section 4 concludes the paper.

2. Materials and Methods

Historical travel time series are extracted from vehicle trajectory data and include a large number of sample points. Each sample trajectory consists of a vehicle ID, a time stamp, longitude, latitude, speed, etc. To attain the travel time of a specific trip, road matching of vehicle trajectory data is performed to specific routes using the method proposed by Li et al. [34]. By calculating the difference in the time stamp of the first and last sample points of origin and destination of a given trip, the set of travel times for all trips is established. Then, based on a predefined departure time interval, the travel times of all trips that fall into the departure time interval are averaged to generate travel time series.

For any travel time series with the same routes, we employ the *RCMSE* algorithm [29] to calculate their multiscale entropy values and evaluate the complexity. Then, travel time predictability is defined, and the relationship between the upper bound of travel time predictability and the entropy of the historical travel time series is presented.

2.1. Entropy of Travel Time Series

Multiscale entropy of the travel time series is measured by the refined composite multiscale entropy (*RCMSE*) algorithm.

Let $X = \{X_1, X_2, \ldots, X_N\}$ denote a travel time series.

Step 1. Construct m-dimensional vectors X_i^m by using Equation (1).

$$X_i^m = \{X_i, X_{i+1}, \ldots, X_{i+m-1}\}, \ 1 \leq i \leq N - m. \tag{1}$$

Step 2. Calculate the Euclidean distance d_{ij}^m between any two vectors X_i^m and X_j^m by using Equation (2).

$$d_{ij}^m = \|X_i^m - X_j^m\|_\infty,\ 1 \le i, j \le N - m,\ i \ne j. \tag{2}$$

Step 3. Let r be the tolerance level. If $d_{ij}^m \le r$, X_i^m and X_j^m are called an m-dimensional matched vector pair. n^m represents the total number of m-dimensional matched vector pairs. Similarly, n^{m+1} is the total number of $(m+1)$-dimensional matched vector pairs.

Step 4. The sample entropy (*SampEn*) is defined by Equation (3).

$$SampEn(X, m, r) = -ln\frac{n^{m+1}}{n^m}. \tag{3}$$

Step 5. Let $y_k^\tau = \{y_{k,1}^\tau, y_{k,2}^\tau, \ldots, y_{k,p}^\tau\}$ be the k-th coarse-grained time series of X defined in Equation (4), where p is the length of the coarse-grained time series, and τ is a scale factor. To obtain y_k^τ, the original time series X is segmented into N/τ coarse-grained series with each segment with a length τ. The j-th element of the k-th coarse-grained time series $y_{k,j}^\tau$ is the mean value of each segment τ of the original time series X.

$$y_{k,j}^\tau = \frac{1}{\tau}\sum_{i=(j-1)\tau+k}^{j\tau+k-1} X_i,\ 1 \le j \le \frac{N}{\tau},\ 1 \le k \le \tau. \tag{4}$$

Step 6. Classical multiscale entropy, $MSE(X, \tau, m, r)$, is defined by Equation (5).

$$MSE(X, \tau, m, r) = SampEn(y_1^\tau, m, r). \tag{5}$$

Step 7. *RCMSE* is defined in Equation (6), where $n_{k,\tau}^m$ is the total number of m-dimensional matched vector pairs in the k-th coarse-grained time series with a length of τ.

$$RCMSE(X, \tau, m, r) = -ln\frac{\sum_{k=1}^\tau n_{k,\tau}^{m+1}}{\sum_{k=1}^\tau n_{k,\tau}^m}. \tag{6}$$

Compared with *SampEn* and *MSE* which are more likely to induce undefined entropy, the *RCMSE* algorithm can estimate entropy more accurately. In Equation (7), the true entropy $S(X)$ of the travel time series X is denoted by $RCMSE(X, \tau, m, r)$. $S(X)$ is roughly equal to, with time scale τ, the negative logarithm of the mean of the conditional probability of new patterns (i.e., the distance between vectors is greater than r) when the dimension of the pattern changes (i.e., m to $m+1$). $S(X)$ describes the degree of irregularity of the travel time series at different time scales and is proportional to the complexity of the travel time series. Based on Equation (7), the true entropy of the travel time series with different time scales can be achieved.

$$S(X) = RCMSE(X, \tau, m, r). \tag{7}$$

2.2. Travel Time Predictability

Based on the historical travel time series, the predictability of travel time is defined as the probability Π that an algorithm can correctly predict future travel time. Again, $X = \{X_1, X_2, \ldots X_N\}$ represents a historical travel time series, φ is the actual travel time of the $(N+1)^{th}$, $\hat{\varphi}$ is the expected value, and φ_a is the estimated value based on model α. Let π be the probability of $\varphi = \hat{\varphi}$ with a given historical travel time series X. Equation (8) shows that π is the random value of distribution of the subsequent travel time and it is an upper bound of the probability distribution of predictive values. In other words, any prediction based on historical series X cannot do better than the one having the true travel time being equal to the expected value, $\varphi = \hat{\varphi}$.

$$\pi = P(\varphi = \hat{\varphi}|X)$$

$$= \sup_x \{P(\varphi_a = x|X)\}$$
$$\geq \mathrm{P}(\varphi = \varphi_a|X). \quad (8)$$

The definition of predictability Π for a travel time series with a length of N is given by Equation (9), where $P(X)$ is the probability of observing a particular historical travel time series X. $\sum \pi P(X)$ presents the best success rate to predict the $(N+1)^{th}$ travel time based on X. Π can be viewed as the averaged predictability (Song et al., 2010) of a historical travel time series.

$$\Pi = \lim_{N\to\infty} \frac{1}{n} \sum_i^N \sum \pi P(X). \quad (9)$$

Next, we relate entropy $S(X)$ to predictability Π to explore the upper bound of predictability Π^{max}. Based on Fano's inequality [24], the relationship between entropy and predictability is shown in Equation (10), which indicates that the complexity of X is less than or equal to the sum of the complexity of successfully predicting $S(\Pi)$ and the complexity of failing to predict $(1-\Pi)\log_2(n-1)$, where n is the number of values of X. Travel time is defined in second and n denotes the total time (seconds) of X. The equality in Equation (10) holds up when Π meets the maximum value Π^{max}. $S(\Pi)$ is presented in Equation (11) and its relationship with the upper bound of travel time predictability is presented in Equation (12). Based on the known $S(X)$ in Equation (7), we can traverse from 0 to 1 to achieve the optimal solution of Π^{max} with a given accuracy target.

$$S(X) \leq S(\Pi) + (1-\Pi)\log_2(n-1) \quad (10)$$

$$S(\Pi) = -\Pi \log_2 \Pi - (1-\Pi)\log_2(1-\Pi) \quad (11)$$

$$S(X) = -\Pi^{max}\log_2\Pi^{max} - (1-\Pi^{max})\log_2(1-\Pi^{max}) + (1-\Pi^{max})\log_2(n-1). \quad (12)$$

3. Experiments

3.1. Study Area and Data

An express road section (about 6.74 km) in Shanghai, China is selected as the study area. As shown in Figure 1, the selected route is a traffic corridor with heavy traffic and is a part of the express road system of Shanghai represented by gray lines. Since this area includes the closed road segments, continuous traffic flow will not be interrupted by an intersection's delay, and the fluctuations of travel time coincide with traffic patterns.

Taxi cab trajectory data associated with the selected route in April 2015 are extracted as the real travel time series. Note that we only use the trips occupied by passengers and discard the patrolling ones (a state of seeking clients in the street). Each trip consists of an origin, a destination, and a route. The total number of passenger trips is 20,430, and averages about 29 per hour, which is sufficient to represent the dynamics of travel time. In addition to the route length, the travel time of a taxi cab trajectory may also be affected by heterogeneous driving behavior, therefore, the distribution of travel time derived from individual trips could be rather complex [6,30,32]. Instead, the travel time estimated from multiple vehicles can well reflect the average trend of travel time and hence more effectively characterize traffic patterns than individual mobility. We use a 5 min time interval to the average travel time of travel cases to obtain a 5 min travel time series from taxi cab trajectory data in April 2015. Figure 2 shows the 5 min travel time series with 8,640 sample GPS points. Let $X = \{X_i | 0 < i < N,\ N = 8640\}$ depict the 5 min travel time series, where X_i is the i-th sample point, and N is the number of points of X.

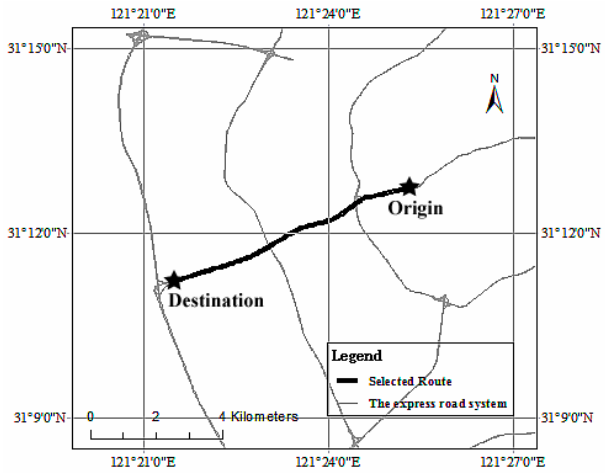

Figure 1. The selected express road section.

Given that the key of this paper is to analyze travel time predictability from the aspect of prediction, the average taxi mobility trends are presented by the 5 min travel time series and have nothing to do with individual taxi behavior. The analysis and evaluation of entropy and predictability are given below.

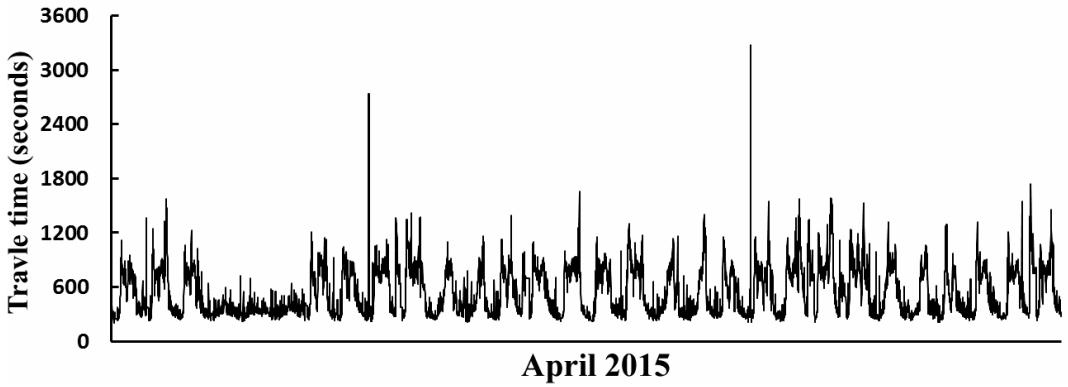

Figure 2. The 5 min travel time series from taxi trajectory data in April 2015.

3.2. Entropy and Predictability

To evaluate the complexity of travel time series, the daily entropy, named S^d, and the weekly entropy, named S^w, are calculated with 24-hour subseries, i.e., 288 consecutive points, and 168 h (7 days) subseries, i.e., 2016 consecutive points, respectively, from the 5 min travel time series. We set the difference of adjacent subseries is 1 h (12 points) to obtain many subseries. Then, we let $S^d = \{S_j^d(X') \Big| X' = \{X_{j\times 12}, X_{j\times 12+1}, \ldots, X_{j\times 12+288-1}\},\ 0 < j < 696\}$, where X' is a subset of X with 288 consecutive points, and we let $S^w = \{S_j^w(X'') \Big| X'' = \{X_{j\times 12}, X_{j\times 12+1}, \ldots, X_{j\times 12+2016-1}\},\ 0 < j < 552\}$, where X'' is a subset of X with 2016 consecutive points.

Let scale factor $\tau = 1$, tolerance $r = 0.1\sigma$, and dimension $m = 2$, where σ is the standard deviation of the 5 min travel time series. The statistical results of values of S^d and S^w are shown in Figure 3. It can be seen that the values of S^d are scattered in the range of 0.6–3.4 and the values of S^w are compact in the range of 1.6–2.3. The remarkable difference between S^d and S^w means that the complexity of daily travel time series tends to change frequently; by contrast, the complexity of weekly travel time series is stable. S^w peaks at about 1.7, indicating that, on average, the probability of new 2-dimension ($m = 2$) patterns in a weekly travel time series is $e^{-1.7} \approx 0.183$. S^d peaks at about 1.2, and the probability

of new patterns in a daily travel time series is $e^{-1.2} \approx 0.301$, indicating that weekly travel time series with greater complexity have lower probability of new patterns than relatively simple daily travel time series.

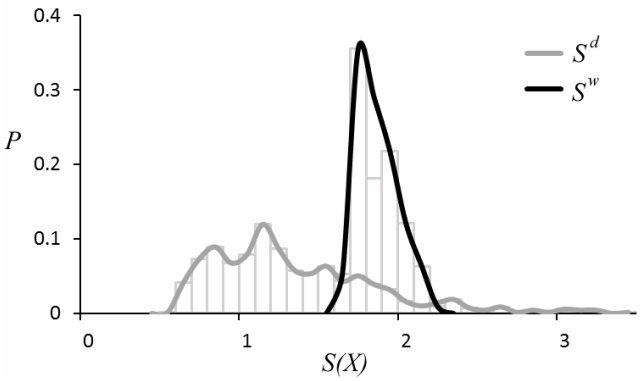

Figure 3. The entropy in the weekly travel time series and the daily travel time series.

Travel time predictability is the probability of an accurate prediction, which is determined by the complexity (entropy) and the range of the travel time series. In our experiments, we set the accuracy of 0.001 to calculate the upper bound of travel time predictability Π^{max} by Equation (12). Therefore, the optimal (maximum) Π^{max} can be achieved by traversing in the range of 0.001–0.999.

The statistical results of the upper bound of travel time predictability in weekly travel time series Π^w and daily travel time series Π^d are shown in Figure 4. Since travel time predictability is influenced by entropy and the range of series, weekly travel time series with higher entropy have lower predictability, peaking at 0.95, and daily travel time series with lower entropy have higher predictability, peaking at 0.99. This demonstrates that the more complex the travel time series is, the less predictability it is, in other words, the less likely to correctly predict it.

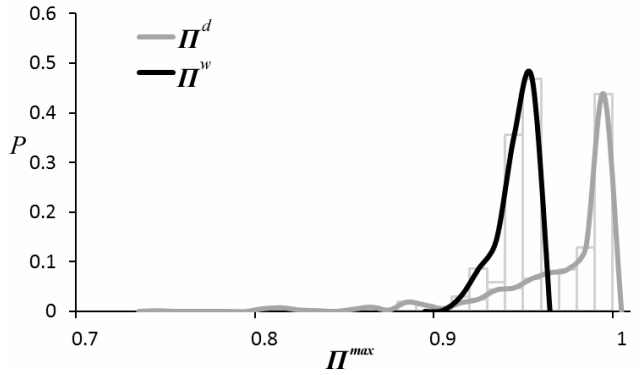

Figure 4. The upper bound of travel time predictability in the weekly travel time series and the daily travel time series.

3.3. Analysis and Discussion

As formulated in Equations (7) and (12), the features of travel time predictability are influenced by three key factors, i.e., scale factor τ, tolerance r, and series length. In this subsection, we analyze the effectiveness of these factors to the predictability of travel time series, and discuss the features and trends of Π^{max}. The validity of the proposed travel time predictability is verified by comparing Π^{max} and the prediction results of future travel time from two typical prediction models, *ARIMA* and *BPNN*.

3.3.1. Time Scales

The scale factor τ is the key parameter of *RCMSE*. It can be used to analyze the complexity and predictability of travel time series in multiple time scales. Figure 5 shows the entropy and predictability of 5 min travel time series with time scales of 1–20. The entropy is calculated by Equation (7) with $r = 0.1\sigma$, and $m = 2$. Π^{max} is calculated by Equation (12). As τ increases, entropy rises and predictability falls. There are more "new patterns" in the travel time series of longer time scales. The complexity of the travel time series of longer time scales is greater than those of shorter time scales, and the travel time series of longer time scales are more difficult to correctly predict.

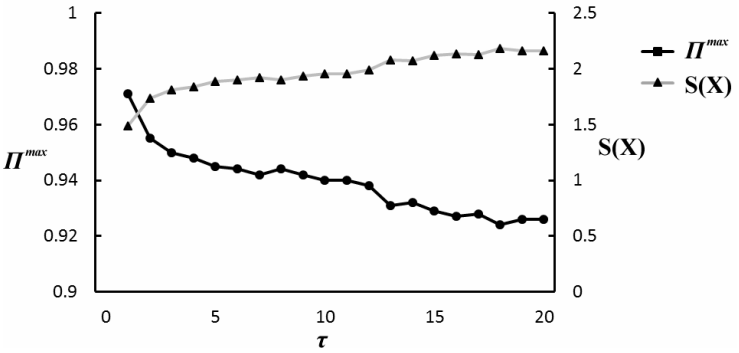

Figure 5. Entropy and predictability of 5 min travel time series with scale factor of 1–20.

3.3.2. Tolerance

Tolerance r, is a key factor in evaluating the complexity of travel time series and constrains the contributions of travel time fluctuations to the complexity. We attempt to evaluate the effectiveness of r in six time scales, i.e., τ = 2, 4, 6, 8, 10, and 12. Figures 6 and 7 show the changing trends of entropy and travel time predictability of the 5 min travel time series with r of 0.01σ to 0.16σ, respectively. Since the travel time predictability of six time scales reaches the maximum value 0.999 when r is equal to 0.16σ, the test range of r is 0.01σ to 0.16σ. In Figure 6, with the increase in r, the entropy gradually becomes lower because the value gap between travel times less than r is not concerned. To six time scales, in addition to τ = 2, other values of entropy are hard to distinguish at a lower r, and ordered values of entropy can be found at higher r values (about 0.06σ to 0.16σ).

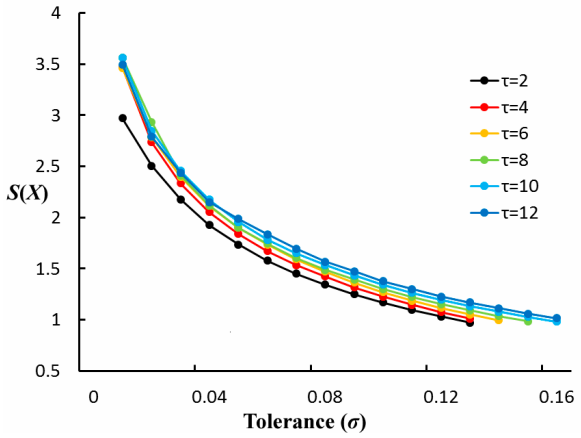

Figure 6. The effectiveness of r to entropy.

Figure 7 shows Π^{max} of the 5 min travel time series with six time scales. There is a negative correlation between Π^{max} and S(X). The higher the Π^{max} is and the lower the S(X) is in shorter time

scales, i.e., τ = 2, the lower the Π^{max} is and the higher the S(X) is in longer time scales. At the same tolerance level, travel time series with a lower r value are easier to predict than those with a higher r. With the expansion of r, Π^{max} gradually increases to 0.999. Π^{max} is limited by r. Obviously, the higher r is, the higher the tolerance is to predictive error, the greater the Π^{max} is, and the more accurate the prediction is.

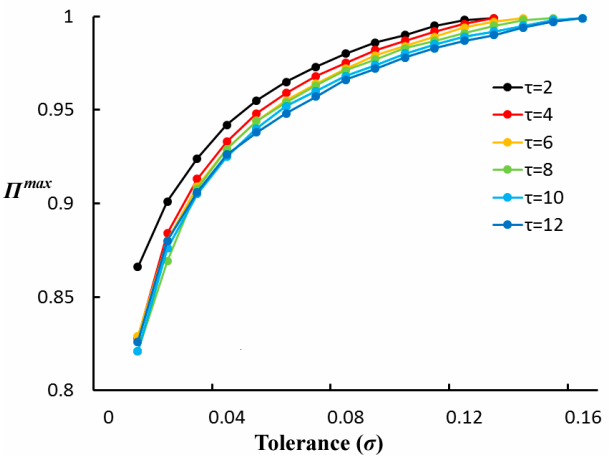

Figure 7. The effectiveness of r to travel time predictability.

For the possibility of perfect theoretical prediction, Figure 8 shows the tolerance of perfect prediction in multiple time scales. The ranges of τ are from 1 to 20. Black line represents the trends of τ with a Π^{max} of 0.999. For example, the next travel time in $\tau = 6$ can be accurately predicted with $r = 0.14\sigma$ by the appropriate prediction model. The growth trend of r indicates that a higher τ is more difficult to predict, and their perfect prediction needs greater tolerance ranges.

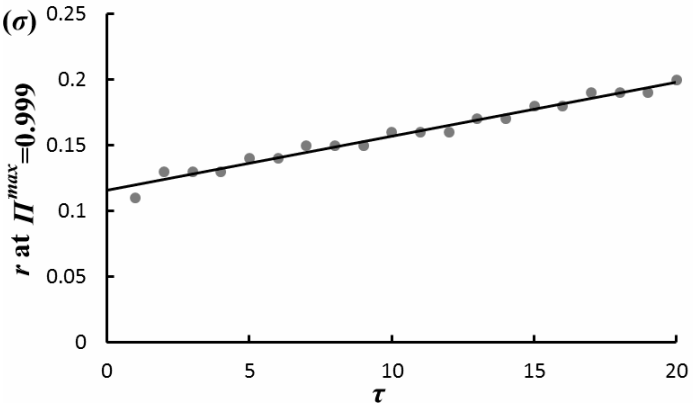

Figure 8. The effectiveness of r to perfect prediction.

3.3.3. Series Length

Next, we analyze the influence of series length on entropy and predictability. Figure 9 shows the entropy of travel time series in six time scales with different series length. The series length of a one-day 5 min travel time series is 288, and so on. Meanwhile, $r = 0.1\sigma$, and $m = 2$. It can be seen that these higher entropy are in 2 day or 3 day travel time series, and the more stable trends are in the >14 day (i.e., two weeks) travel time series. We can think that entropy of a >14 day travel time series is roughly independent of series length. Table 1 shows that the statistics of entropy to support these findings, where $\overline{S(X)}$ is the average value of entropy of all travel time series, and sd_S is the standard deviation of all entropy. The sd_S of >14 days is much lower than those of <14 days, and the most stable

series is the >14 day travel time series with $\tau = 2$ and $\tau = 4$. Therefore, we can demonstrate that the complexity of the >14 day travel time series is stable and is independent of series length.

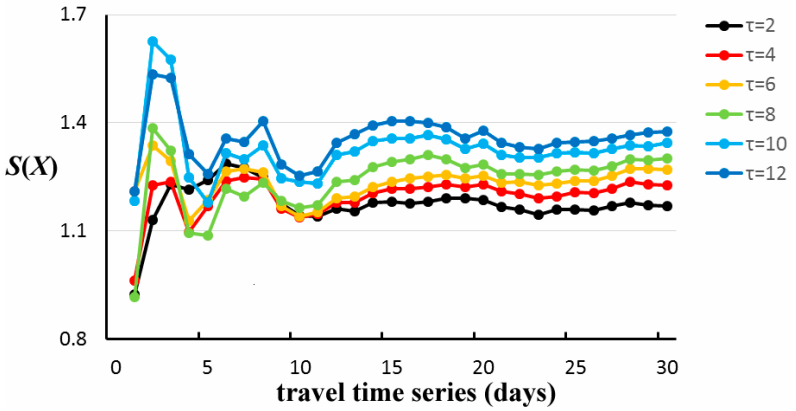

Figure 9. Entropy of travel time series with different series lengths.

Similarly, Figure 10 demonstrates the smallest value of predictability of 2 day or 3 day travel time series and the stationarity and independence of predictability of the >14 day travel time series. In Table 1, $\overline{\Pi^{max}}$ denotes the average value of travel time predictability, and $sd_{\Pi^{max}}$ denotes the standard deviation of travel time predictability. Great differences between the $sd_{\Pi^{max}}$ of the >14 day travel time series and the <14 day travel time series present the stable predictability of the >14 day travel time series. In addition, we can demonstrate that the most stable predictability is in the >14 day travel time series with $\tau = 2$ and $\tau = 4$.

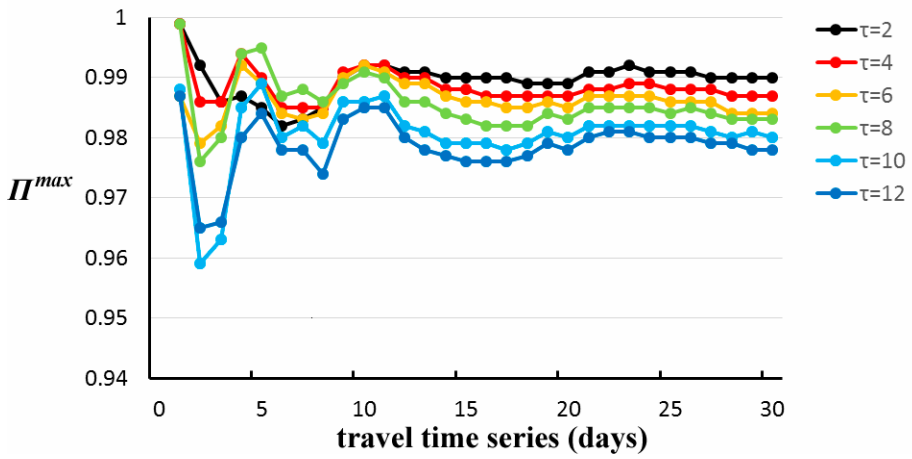

Figure 10. Travel time predictability of travel time series with different series lengths.

Table 1. The statistics of entropy and travel time predictability.

τ	$\overline{S(X)}$	sd_s		$\overline{\Pi^{max}}$	$sd_{\Pi^{max}}$	
		$<$ 14 Days	$\geq$ 14 Days		$<$ 14 Days	$\geq$ 14 Days
2	1.1752	0.0896	0.0126	0.9896	0.0045	0.0008
4	1.1966	0.0755	0.0125	0.9885	0.0040	0.0007
6	1.2334	0.0623	0.0148	0.9863	0.0040	0.0011
8	1.2415	0.1108	0.0167	0.9856	0.0058	0.0011
10	1.3261	0.1311	0.0193	0.9805	0.0089	0.0013
12	1.3571	0.0953	0.0242	0.9786	0.0066	0.0016

3.3.4. The Validity of Travel Time Predictability

To validate the travel time predictability, two prediction models, i.e., AutoRegressive Integrated Moving Average (*ARIMA*) [35], and Back Propagation Neuro Networks (*BPNN*) [36], are employed to predict future travel time.

The *ARIMA* model is a method for time series analysis and prediction. Since travel time series have obvious fluctuation differences between weekdays and weekends, we use the seasonal *ARIMA* (*SARIMA*) model, denoted $ARIMA(p,d,q)(P,D,Q)_m$, to predict future travel time, where p is the order of the autoregressive (*AR*) part, q is the order of the moving average (*MA*) part, d is the degree of difference for reducing the non-stationarity of time series, m is the number of periods per season, and P, D, Q refer to the *AR*, differencing, and *MA* terms for the seasonal part of the *ARIMA* model. Due to the stationary and weekly change period of travel time series in our experiments, we set $d = 0$, $D = 0$, and $m = 7$. By testing the autocorrelation function (*ACF*) and the partial autocorrelation function (*PACF*) of complete and seasonal part travel time series, we set $p = 3$, $q = 1$, $P = 1$, and $Q = 2$. Then, we use $ARIMA(3,0,1)(1,0,2)_7$ to predict future travel time in the selected route.

As a neuro network method, the *BPNN* model includes an input layer, a hidden layer, and an output layer. It can learn and store large amounts of input–output mapping by model training to represent and predict the dynamic and non-linear processes. In our experiments, the *BPNN* model has three inputs, i.e., the date, the time of day, and the day of week, and one output, i.e., the travel time, the number of nodes in the hidden layer is 7, the learning rate (η) is 0.9, and the momentum factor (α) is 0.7.

Figure 11 shows the errors of travel time prediction with *ARIMA* and *BPNN* models in 5 min travel time series. We set $r = 0.1\sigma$ (about 22 s), $m = 2$, and $\tau = 2$. We predict 50 times with *ARIMA* and *BPNN*, respectively, and let e be the absolute value of the difference between predictive value and actual value. The dashed line indicates the tolerance $r = 0.1\sigma$. It can be seen that most dots are below it. The statistical results of travel time prediction of the 5 min travel time series are shown in Table 2. $\overline{\Pi^{max}}$ denotes the average predictability of 50 predictions in the 5 min travel time series. If e is less than r (below the dashed line of Figure 11), it is a successful prediction. The number of successful predictions is 40 with *ARIMA*, and 41 with *BPNN*. Compared with $\overline{\Pi^{max}}$ of 0.952, the success rates of prediction are lower, while their average errors, 13.46 and 12.42, are lower than tolerance r, 22.

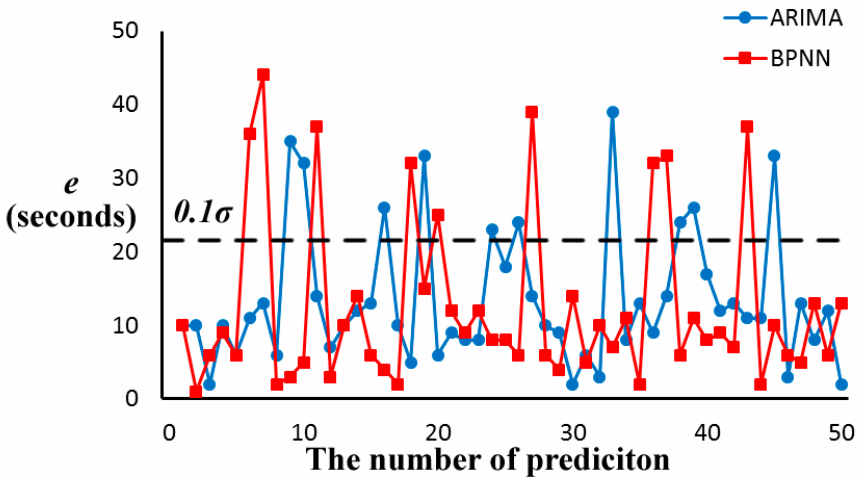

Figure 11. The errors of travel time prediction in the 5 min travel time series.

Table 2. The statistics of prediction with the 5 min travel time series.

Prediction Model	Number of Prediction	Number of Success	Success Rate	Average Error (s)	$\overline{\Pi^{max}}$
ARIMA	50	38	90%	13.46	0.952
BPNN	50	40	91%	12.42	

To comprehensively evaluate the relationships between travel time predictability and prediction results, two group comparisons were conducted. Figure 12 shows the comparison results between travel time predictability and prediction results in the 5 min travel time series with different series lengths of 1–30 days. The average prediction results of 100 experiments of each travel time series are presented: $r = 0.1\sigma$, $m = 2$, and $\tau = 2$. Let SR_{ARIMA} be the success rate of travel time prediction by *ARIMA*, and let SR_{BPNN} be the success rate by *BPNN*. There is a significant statistical difference between SR_{ARIMA} (or SR_{BPNN}) and $\overline{\Pi^{max}}$. It indicates that the actual performance (SR_{ARIMA} and SR_{BPNN}) of *ARIMA* and *BPNN* lags far behind the theoretical optimal value of success rate of travel time prediction, and the performance of travel time prediction is still great room for improvement. Note that the change trends of SR_{ARIMA} or SR_{BPNN} are basically consistent, as shown in Table 3, with the standard deviation (*sd*) of the travel time predictability, and the predicted results of *ARIMA* and *BPNN* with shorter, 14-day series lengths have obvious higher values than the longer, 14-day series length, which indicates that the accuracy of prediction is affected by the complexity of the travel time series and demonstrates the validity of travel time predictability.

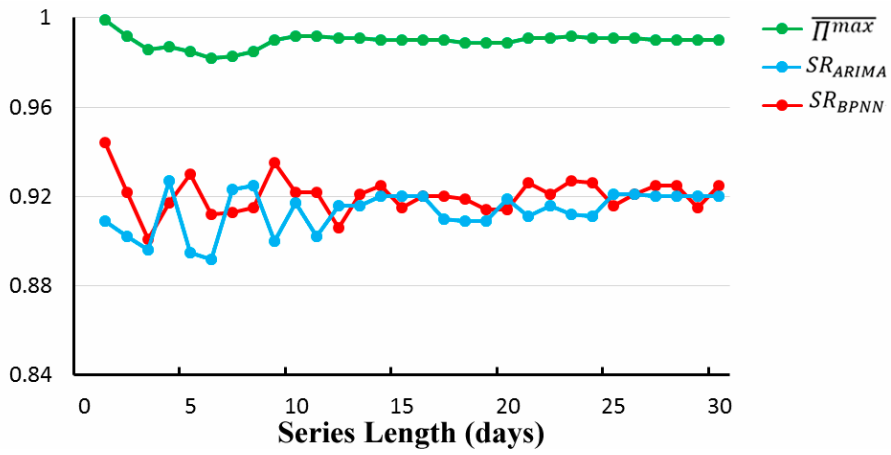

Figure 12. The comparisons between travel time predictability and prediction results in the 5 min travel time series with different series lengths.

Table 3. The standard deviation (*sd*) of travel time predictability and the predicted results with different series lengths.

Series Length (Days)	$\overline{\Pi^{max}}$	SR_{ARIMA}	SR_{BPNN}
$<$ 14 days	0.0052	0.0339	0.0378
$\geq$ 14 days	0.0009	0.0142	0.0137

The same situation occurs in Figure 13. We compare $\overline{\Pi^{max}}$ and the prediction results in different time scales of 1–20 with $r = 0.1\sigma$, and $m = 2$. With the increase of time scales, travel time predictability and the success rates of two prediction models decline synchronously.

The proposed travel time predictability is a valid measurement of travel time series for correct prediction, which provides an achievable target to the development of travel time prediction methods and contribute to a differentiated scheme of travel time prediction.

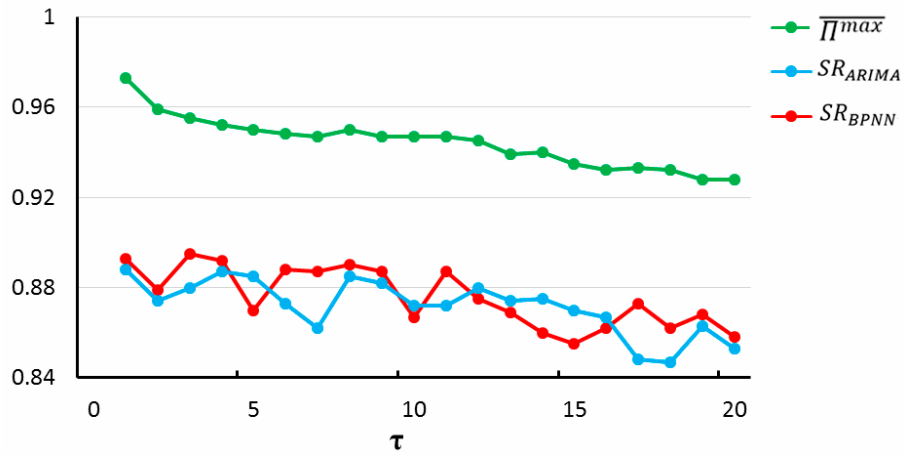

Figure 13. Relationships between travel time predictability and prediction results in travel time series with different time scales.

4. Discussion and Conclusions

This paper defines travel time predictability as the probability of correctly predicting future travel times based upon historical travel time series and develops an entropy-based approach to measure the upper bound of travel time predictability. Multiscale entropy of travel time series is calculated to evaluate its complexity. The upper bound of travel time predictability is found to be related to entropy. Travel time predictability expresses the characteristics of travel time series itself and is an expected value of data-based prediction performance.

A case study in an express section road in Shanghai, China is designed. The data source is a large amount of taxi cab trajectory data collected in April 2015. By analyzing the effectiveness of the time scales and tolerance to entropy and travel time predictability, we demonstrate that time scales and tolerance are positively related to the entropy and negative related to travel time predictability. In addition, we reveal the higher value of entropy and the lower predictability of 2 day or 3 day travel time series and the more stable values of >14 day travel time series. Finally, two prediction models, *ARIMA* and *BPNN*, are employed to predict travel time based on historical travel time series and to examine the validity and reliability of travel time predictability. Though travel time predictability is independent of the prediction method, it can aid the development of travel time prediction methods and contribute to a differentiated scheme for travel time prediction in diverse traffic environment.

Future efforts may be pursued in two directions. First, the comprehensive investigation and verification of travel time predictability should begin in a larger network with multiple data sources to show the capability of capturing the entropy, which contributes to deeper traffic knowledge discovery and differentiation traffic police formulation. Second, the scope of predictability should be extended and the possibility of applying predictability to other types of time series may be surveyed.

Acknowledgments: This research was sponsored by the National Natural Science Foundation of China (Grant No. 41271441). The authors also would like to show their gratitude to 3 anonymous reviewers for their constructive comments that greatly improved the manuscript.

Author Contributions: Tao Xu developed the methodology and composed the manuscript; Xianrui Xu conducted experiments and analyzed the results; Yujie Hu implemented the prediction methods; Xiang Li proposed the initial idea and completed the literature review.

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

References

1. Wu, C.H.; Ho, J.M.; Lee, D.T. Travel-time prediction with support vector regression. *IEEE Trans. Intell. Transp. Syst.* **2005**, *5*, 276–281. [CrossRef]

2. Mori, U.; Mendiburu, A.; Álvarez, M.; Lozano, J.A. A review of travel time estimation and forecasting for Advanced Traveller Information Systems. *Transportmetr. A Transp. Sci.* **2015**, *11*, 1–39. [CrossRef]
3. Oh, S.; Byon, Y.J.; Jang, K.; Yeo, H. Short-Term Travel-Time Prediction on Highway: A Review of the Data-Driven Approach. *Transp. Rev.* **2015**, *35*, 4–32. [CrossRef]
4. Vlahogianni, E.I.; Karlaftis, M.G.; Golias, J.C. Short-term traffic forecasting: Where we are and where we're going. *Transp. Res. Part C Emerg. Technol.* **2014**, *43*, 3–19. [CrossRef]
5. Lin, H.E.; Zito, R.; Taylor, M. A review of travel-time prediction in transport and logistics. *Proc. East. Asia Soc. Transp. Stud.* **2005**, *5*, 1433–1448.
6. Van Lint, J.W.C.; van Hinsbergen, C.P.I. Short-Term Traffic and Travel Time Prediction Models. *Artif. Intell. Appl. Crit. Trans. Issues* **2012**, *22*, 22–41.
7. Vlahogianni, E.I.; Golias, J.C.; Karlaftis, M.G. Short-term traffic forecasting: Overview of objectives and methods. *Transp. Rev.* **2004**, *24*, 533–557. [CrossRef]
8. Abrantes, P.A.L.; Wardman, M.R. Meta-analysis of UK values of travel time: An update. *Transp. Res. Part A Policy Pract.* **2011**, *45*, 1–17. [CrossRef]
9. Jara-Diaz, S.R. *Transport Economic Theory*; Emerald Group Publishing Limited: Bingley, UK, 2007; pp. 11–49.
10. Li, Z.; Hensher, D.A.; Rose, J.M. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transp. Res. Part E Log. Transp. Rev.* **2010**, *46*, 384–403. [CrossRef]
11. Van Lint, J.W.C.; Zuylen, H.J.V.; Tu, H. Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transp. Res. Part A Policy Pract.* **2008**, *42*, 258–277. [CrossRef]
12. Shires, J.D.; de Jong, G.C. An international meta-analysis of values of travel time savings. *Eval. Program Plan.* **2009**, *32*, 315–325. [CrossRef] [PubMed]
13. Wardman, M.; Batley, R. Travel time reliability: A review of late time valuations, elasticities and demand impacts in the passenger rail market in Great Britain. *Transportation* **2014**, *41*, 1041–1069. [CrossRef]
14. Noland, R.B.; Polak, J.W. Travel time variability: A review of theoretical and empirical issues. *Transp. Rev.* **2002**, *22*, 39–54. [CrossRef]
15. Carrion, C.; Levinson, D. Value of travel time reliability: A review of current evidence. *Transp. Res. Part A Policy Pract.* **2012**, *46*, 720–741. [CrossRef]
16. Rietveld, P.; Bruinsma, F.R.; Vuuren, D.V. Coping with unreliability in public transport chains: A case study for Netherlands. *Transp. Res. Part A Policy Pract.* **2001**, *35*, 539–559. [CrossRef]
17. Yue, Y.; Yeh, A.G.O.; Zhuang, Y. Prediction time horizon and effectiveness of real-time data on short-term traffic predictability. *Proc. Intell. Transp. Syst. Conf.* **2007**, 962–967.
18. Foell, S.; Phithakkitnukoon, S.; Kortuem, G.; Veloso, M. Predictability of Public Transport Usage: A Study of Bus Rides in Lisbon, Portugal. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2955–2960. [CrossRef]
19. Siddle, D. Travel Time Predictability. Available online: https://trid.trb.org/view.aspx?id=1323043 (accessed on 11 April 2017).
20. Wang, J.; Mao, Y.; Li, J.; Xiong, Z.; Wang, W.X. Predictability of road traffic and congestion in urban areas. *PLoS ONE* **2015**, *10*, e0121825. [CrossRef] [PubMed]
21. Du, Y.; Chai, Y.W.; Yang, J.W.; Liang, J.H.; Lan, J.H. Predictability of Resident activity in Beijing Based on GPS Data. *Geogr. Geo Inf. Sci.* **2015**, *31*, 47–51.
22. Song, C.; Qu, Z.; Blumm, N.; Barabási, A.L. Limits of predictability in human mobility. *Science* **2010**, *327*, 1018–1021. [CrossRef] [PubMed]
23. Shannon, C.E. A Mathematical Theory of Communication: The Bell System Technical Journal. *Bell Syst. Tech. J.* **1948**, *27*, 3–55. [CrossRef]
24. Fano, R.M.; Hawkins, D. Transmission of Information: A Statistical Theory of Communications. *Am. J. Phys.* **1961**, *29*, 793–794. [CrossRef]
25. Kontoyiannis, I.; Algoet, P.H.; Suhov, Y.M.; Wyner, A.J. Nonparametric entropy estimation for stationary processes and random fields, with applications to English text. *IEEE Trans. Inf. Theory* **2007**, *44*, 1319–1327. [CrossRef]
26. Wyner, A.D.; Ziv, J. The sliding-window lempel-ziv algorithm is asymptotically optimal. *Proc. IEEE* **1994**, *82*, 872–877. [CrossRef]
27. Xie, X.X.; Li, S.; Zhang, C.L.; Li, J.K. Study on the application of Lempel-Ziv complexity in the nonlinear detecting. *Complex Syst. Complex. Sci.* **2005**, *2*, 61–66.

28. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *92*, 705–708. [CrossRef] [PubMed]
29. Wu, S.D.; Wu, C.W.; Lin, S.G.; Lee, K.Y.; Peng, C.K. Analysis of complex time series using refined composite multiscale entropy. *Phys. Lett. A* **2014**, *378*, 1369–1374. [CrossRef]
30. González, M.C.; Hidalgo, C.A. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779. [CrossRef] [PubMed]
31. Pappalardo, L.; Simini, F.; Rinzivillo, S.; Pedreschi, D.; Giannotti, F.; Barabási, A.L. Returners and explorers dichotomy in human mobility. *Nat. Commun.* **2015**, *6*. [CrossRef] [PubMed]
32. Pappalardo, L.; Rinzivillo, S.; Qu, Z.; Pedreschi, D.; Giannotti, F. Understanding the patterns of car travel. *Eur. Phys. J. Spec. Top.* **2013**, *215*, 61–73. [CrossRef]
33. Hu, Y.; Miller, H.J.; Li, X. Detecting and analyzing mobility hotspots using surface networks. *Trans. GIS* **2014**, *18*, 911–935. [CrossRef]
34. Li, X.J.; Li, X.; Tang, D.; Xu, X. Deriving features of traffic flow around an intersection from trajectories of vehicles. In Proceedings of the 2010 International Conference on Geoinformatics: Giscience in Change, Geoinformatics, Beijing, China, 18–20 June 2010; pp. 1–5.
35. Box, G.E.P.; Jenkins, G.M. Time series analysis: Forecasting and control. *J. Oper. Res. Soc.* **1971**, *22*, 199–201.
36. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Cogn. Model.* **1988**, *5*, 1. [CrossRef]