Potential Predictability of Vehicles' Visiting Duration in Different Areas For Large Scale Urban Environment

Wenyu Ren, Yong Li, Siyu Chen, Depeng Jin, Li Su
State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
Email: liyong07@mail.tsinghua.edu.cn

***Abstract*—The newly emerged vehicular communication network is seen as a key technology for solving the increasingly serious vehicular traffic congestion as well as improving road safety, and applications of vehicular networks are also emerging at the same time. Predicting vehicular visiting time is vital to both solving the vehicular system problem and building efficient vehicular networking. It is an open and unsolved problem that how much the vehicular staying duration of visits to different areas can be predicted. In this paper, we use real vehicular traces in Beijing and Shanghai and an area partition model to explore the limit of predictability of the vehicular visiting time in different areas in large scale cities and analyze the effect of different precision and slot time on predictability. We conclude that using a proper time slot is an efficient way of prediction and a higher predictability can be achieved if the requirement for precision is reduced. Among all the cases we study, we find the highest potential predictability of 76.3% in Beijing and 82.5% in Shanghai in the case with smallest slot time and lowest requirement for precision.**

I. Introduction

Urban vehicular traffic congestion is an increasingly serious problem which is significantly affecting many aspects of the quality of metropolis life all around of the world [1]. Local governments of large cities are continuing to test the capacities of city roads, and make new adjustment and planning to the existing transportation infrastructure, which attracts lots of financial spending on increasing the capacities of the cities' transportation [2]. Scientific traffic engineering, which achieves efficient resource management in the networks of road and transportation system, becomes a hot research topic attracting broad interests from many communities.

How to better deal with problems of transportation system is very important, and newly emerged vehicular communication network is seen as a key technology for helping relieving the traffic congestion, and at the same time improving road safety, by building intelligent transportation systems [3]. Recently, as more and more vehicles are equipped with devices to provide wireless communication capacities, interests on vehicular communications and networks have grown significantly [3]. Many applications of vehicular networks are also emerging, including automatic collision warning, remote vehicle diagnostics, emergency management and assistance for safe driving, vehicle tracking, automobile high speed Internet access, and multimedia content sharing. In the USA, Federal Communications Commission has allocated 75 MHz of spectrum for dedicated short-range communications in vehicular networks [4], and IEEE is also working on the related standard specifications. Many consortia and standardisation bodies are actively developing technologies and protocols for information transmission between vehicles and roadside unit infrastructure equipments.

It will come as little surprise that the vehicular transportation networks and the vehicular communication networks are simultaneously making a conscious effort toward dealing with transportation problems for urban cities in the intelligent transportation systems. In terms of networks formed by vehicles, which is the transportation system, we need to efficiently use existing road network systems and information like vehicular movements and distributions collected to make impacts on reducing traffic congestion and travel delays, and further on saving energy consumption and improving safety [5]. Thus, effective and accurate real-time predictions of vehicular staying duration in the city area are needed [1] to estimate the vehicular traffic and further predict the congestion events. While on the other hand, as for vehicular communication networks [6], it is hard to maintain a connected and stable network to communicate. Thus, they are usually distributed, self-organized by mobile vehicles, and characterized by very frequent movements and limited communication opportunities in nodes mobility patterns. To this end, the capability of predicting the time of a vehicle's next movement can play a significant role in lots of communication and networking functions from bandwidth reservation to service provisioning.

Combing the above two aspects, we have identified that vehicular staying duration prediction is vital to both solving the vehicular system problem and building efficient vehicular networking. Current, lots of works fall into the area of prediction algorithms [7], [8]. However, how much the vehicular staying duration can be predicted is an open and unsolved problem. In this paper, we consider the problem of the limit of predictability of the vehicular staying duration in different areas in the large scale cities. Specifically, we use intersections of city roads to divide the urban area and get the staying duration of each vehicle in each area from the vehicular traces. The questions we address here is that are there regularities existing to govern the vehicular mobility in terms of the staying time in each area, and how can this regularities influence the predictability of their visiting duration. Our contributions are summarized as follows.

1) We use linear interpolation to preprocess real vehicular mobility traces from two large-scale vehicular datasets in Shanghai and Beijing on which our work is based, and propose a model by selecting hundreds of intersections in Beijing and Shanghai and partitioning the two cities by these intersections' Voronoi cells.
2) We analyze the datasets and find regular time patterns in the mass of data, which indicates the potential predictability in the staying duration of visits to each area.
3) We calculate the entropy and limits of predictability of vehicles' staying duration in each area for both datasets and explore the effect of precision and slot time on the predictability. We conclude that reducing the precision and using proper time slot are both effective ways of raising predictability, and we get the highest potential predictability of 76.3% in Beijing and 82.5% in Shanghai when we use smallest slot time and lowest requirement for precision.

The rest of this paper is organized as follows. After presenting the related work in Section II, we introduce the datasets we used and the preprocessing in Section III. Our model and the problem is described in Section IV. Section V acts as a brief introduction to the our methodology and our calculation results and analysis are shown in Section VI. Finally, we conclude the paper in Section VII.

II. Related Work

Some mobility models for vehicular networks have been proposed by previous works. In [9], Ali *et al*. have made an adaptation of the model MBMM which is aimed at studying human mobility, and presented V-MBMM. Based on real GPS trace data from taxis, META [10] is a model for simulating movement of taxis in a metropolis. Inspired by product-form queuing networks, Mohimani *et al*. designed a mobility model representing a sparse situation for VANETs [11]. Besides these, there are also generators of mobility patterns for VANETs such as MOVE [12] and Citymob [13], which can be used to create various mobility scenarios. Although there are such ready-made models for VANETs existed, they are unsuited and too complicated for our work focused on prediction calculation and analysis. Therefore, we adopt a simple but effective model established on our area partition method.

Prediction is another major research topic in the domain of vehicular networks. There is a concept named lifetime in previous works similar to the staying duration in this paper. In [14], the lifetime is defined as the duration time that two vehicles spending in the communication range of each other, which can be seen as the lifetime of a link. On the other hand, Wan *et al*. in their work [15] used the concept of lifetime of a routing path which consists of several links connected end to end. The stay time in [16] has almost the same meaning to the staying duration in our work. It is worth pointing out that all of these works calculate the duration or time based on the location and speed of the vehicles. In this paper, however, we evaluate the staying duration by vehicles' mobility histories and focus on the fundamental question that to what degree can the staying duration be predicted.

The way we calculate the limits of predictability in our work is previously used by Song *et al*. when dealing with human mobility [17]. Nevertheless, there are three main differences between their work and ours: (1) we explore the predictability in vehicles' staying duration to different areas in this paper, which has entirely distinct properties compared with human mobility; (2) our work is based on real taxi traces gathered by GPS devices, while their data is collected on mobile network carriers; (3) we study the influence of precision and slot time on predictability, which is not included in their analysis.

III. Datasets and Preprocessing

As the foundation of our entire work, vehicular mobility traces in Shanghai and Beijing are of great importance to our calculation and analysis. In order to explain the problem and our model, we first give a general idea of these two large-scale datasets should be given first.

In the *Shanghai* dataset [18] and *Beijing* dataset, the data was recorded by GPS devices deployed on taxis in the city of Shanghai and Beijing and collected through GPRS. The data intervals in different situations and vehicle numbers of the two datasets are shown in Table I. To our knowledge, *Beijing* dataset is so far the largest dataset existed of vehicular traces.

Dataset	Data Interval (second)		Vehicle Number
	passengers onboard	no passenger	
Shanghai	60	15	2,109
Beijing	128	128	27,000

TABLE I
Data interval and vehicle number of *Beijing* and *Shanghai* datasets

Besides the taxi's current position represented by the longitude and latitude coordinates, each report also contains the taxi's ID, instantaneous velocity and heading, and a timestamp. Before we use the traces in the datasets, we first preprocess them to reduce the inconvenience brought by the various data intervals. Some location points are inserted with the method of linear interpolation (LI) and some original points are deleted so that the points have a sampling time of 10 seconds. To validate the traces and our preprocessing, we plot all the preprocessed traces and find that they match the real roads on the city maps pretty well.

IV. Problem and Model Description

Our goal here is to explore regular patterns and measure predictability of vehicle mobility in the traces. This problem

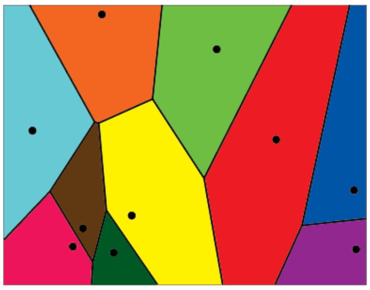


Fig. 1. Illustration of area partition using Voronoi cell: 10 points and their Voronoi cells.

have two aspects: location and time. The location aspect represents the location selection of vehicles, while the time aspect considers the staying duration to each area. In this paper, we focus on the time aspect and discuss our calculation and analysis about the predictability of vehicles' staying duration to different areas in *Beijing* and *Shanghai* datasets.

Because the staying duration is to some extent a reflection of the traffic condition nearby, it should be studied for every area instead of every vehicle. For each area, we assume that it has been visited for *n* times in the past. Denote its time history by $h_n = \{T_1, T_2, \ldots, T_n\}$, where T_k corresponds to the staying duration of the *k*th visit to the area for $1 \leq k \leq n$. With the knowledge of an area's time history h_n, we are curious about the question that what role does randomness play in vehicles' choosing their staying duration in that area and to what degree is the staying duration predictable. To answer this question, we calculate fundamental limits of the predictability in vehicles' staying duration and explore influence of precision and slot time on the potential predictability.

As our work is not focused on model designing, we base our analysis on a simple but effective model. We divide the city of Shanghai and Beijing into 559 and 760 areas respectively using main intersections within coverage of the traces. Each area *i* is defined as the Voronoi cell of the corresponding intersection A_i, which consists of every point whose distance to A_i is less than or equal to its distance to any other intersection [19]. 10 points and the corresponding Voronoi cells are shown in Fig. 1 for a better understanding of the concept.

To further reduce the complexity of our analysis, we then turn each vehicle's complete trace into a sequence consists of items in chronological order. Every item includes an area's index, the moment of entering it and the staying duration in it. Since the preprocessed traces have interval of 10 seconds, we further use the LI method to better estimate the entering moment and staying duration. To illustrate how this LI method works, consider we have two pieces of successive location information of one taxi in the preprocessed traces with the locations l_1 and l_2 recorded at the time points $t_1 < t_2$. Suppose l_1 and l_2 belong to two different areas i_1 and i_2, whose corresponding intersections have locations at A_1 and A_2. With the estimation accuracy set to 1 second, we divide the time between t_1 and t_2 into ten equal parts by $t_{1,1} < t_{1,2} < \cdots < t_{1,9}$ and rename t_1 and t_2 to $t_{1,0}$ and $t_{1,10}$. Hence, $t_{1,0}, t_{1,1}, t_{1,2}, \cdots, t_{1,9}, t_{1,10}$ are 11 increasing successive integers. We calculate the location $l_{1,j}$ at $t_{1,j}$ $(1 \leqslant j \leqslant 9)$ by the following LI

$$l_{1,j} = \frac{t_{1,10} - t_{1,j}}{t_{1,10} - t_{1,0}} \cdot l_1 + \frac{t_{1,j} - t_{1,0}}{t_{1,10} - t_{1,0}} \cdot l_2. \tag{1}$$

Rename l_1 and l_2 to $l_{1,0}$ and $l_{1,10}$. $\exists j_{min}$, $j_{min} \in \{0, 1, \ldots, 10\}$, s.t. $|dis(l_{1,j_{min}}, A_1) - dis(l_{1,j_{min}}, A_2)| = \min_j |dis(l_{1,j}, A_1) - dis(l_{1,j}, A_2)|$, where $dis(A, B)$ is the distance between locations *A* and *B*. The moment of entering area i_2 can be estimated as $t_{1,j_{min}}$. After we estimate the starting time of every visit to every area, we can readily calculate the staying duration by subtraction. Because we add an item to the sequence only when there is a change of visiting area of the vehicle, every successive two items have different area indexes.

In our data, the accuracy of the staying duration is set to 1 second. But sometimes we do not need our prediction to be that precise, hence we introduce a parameter μ to our model, which represents the minimum unit of staying duration under our discussion, in other words, the precision. The higher μ is, the lower the precision is. We base our analysis on five scenarios with μ set to 1, 5, 10, 30 and 60 seconds respectively. To better illustrate this parameter, consider the case with $\mu = 5s$. We call staying duration which is multiple of 5 as legal duration. Then we round each staying duration in the datasets to its nearest legal duration and use the new value in our calculation later. In this case, any predictive algorithm can only predict legal staying duration. Therefore, the larger μ is, the less possible number of values the staying duration could take.

To study the prediction of staying duration in more detail, we bring the concept of time slot into our work. We divide 24 hours of one day into several time slots with equal length and assume that the traffic condition of an area remains almost the same in each time slot for different days. We denote the slot time by τ and set it to 10, 30 and 60 minutes respectively in analyzing the predictability. To compare the difference of predictability before and after we introduce time slot, we also analyze the unslotted case.

V. Methodology

Entropy is a quantity characterizing the uncertainty of a random variable within the scope of information theory [20]. The entropy of a discrete random variable X whose possible values are $\{x_1, \ldots, x_n\}$ is defined as $H(X) \equiv -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$, where $p(X)$ is the probability mass function. Take $n = 2$ for example, if $p(x_1) = 0$ and $p(x_2) = 1$, $H(X) = 0$ and X is entirely predictable for it always gets the value of x_2. However, there are no way for us to predict the next value of X with high accuracy if $p(x_1) = p(x_2) = 0.5$, in which case the entropy is 1, because the two values appear randomly with equal probabilities. In this sense, entropy can tell a lot about predictability. For a time series, the lower its entropy is, the less random it is, and the more predictable the next element is. Fundamentally, entropy can act as an efficient measurement of the degree of predictability [21].

For an area *i*, denote the number of distinct staying durations in its time history by N_i. We define the area's entropy as $S_i \equiv -\sum_{j=1}^{N_i} p_i(j)\log_2 p_i(j)$, where $p_i(j)$ is the historical probability that staying duration *j* appeared at area *i*. Based on the entropy, we calculate the limits of predictability in vehicular staying duration with a method previously used in dealing with human mobility [17]. A brief introduction to the methodology is given first.

For an area with time history h_{n-1}, let $\omega(h_{n-1})$ be the best accuracy that any prediction algorithm can achieve with the knowledge of the history. The predictability in the staying duration of impending *n*'s visit can be defined as

$$\Omega(n) \equiv \sum_{h_{n-1}} P(h_{n-1})\omega(h_{n-1}), \tag{2}$$

where $P(h_{n-1})$ is the probability of observing the particular history h_{n-1}. Based on this, the overall predictability of the area is defined as

$$\Omega \equiv \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \Omega(i). \tag{3}$$

Consider an area with *N* possible staying durations and entropy *S*, it has been proven [17] that the predictability Ω's upper bound Ω_{max} can be calculated by

$$\begin{aligned} S = &-\left[\Omega_{max}\log_2\Omega_{max} + (1-\Omega_{max})\log_2(1-\Omega_{max})\right] \\ &+ (1-\Omega_{max})\log_2(N-1). \end{aligned} \tag{4}$$

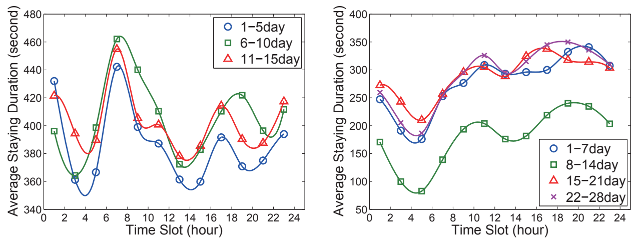
For an area with $\Omega_{max} = 0.6$, 60% of the visits to that area last for a duration that may be predicted, and the staying duration to the rest 40% of the visits appear to be random. In other words, no prediction algorithm in the world can attain an accuracy better than 60% in the long run when predicting staying duration in that area. Hence, Ω_{max} measures the fundamental limit of the area's predictability in staying duration.

VI. Results and Analysis

In this section, we present our calculation and analysis results. First, we analyze the traces in the two datasets and find that there are potential time patterns hidden beneath the surface. We then calculate the entropy and give the limits of predictability in vehicles' staying duration. To explore the effect of slot time and precision on this predictability, we also make comparisons among different cases.

A. Time Pattern

To explore the time pattern in the traces, we study the area staying duration of a single area. An area is randomly selected from the most crowded 50 areas respectively in both cities. We then divide each day into 12 time slots, with every slot covering 2 consecutive hours. After that, we calculate the average staying duration of visits to the selected area in every time slot every 5 days in Beijing and 7 days in Shanghai (see Fig. 2(a) and Fig. 2(b)). In these two figures, we can clearly find that the average staying duration follows the same pattern over different days and the morning and evening peaks can be recognized easily. The traffic condition deteriorates between $8:00 \sim 12:00$ and $16:00 \sim 19:00$ in *Beijing* dataset while between $6:00 \sim 9:00$ and $18:00 \sim 21:00$ in *Shanghai* dataset. From the time pattern we find in the traces, we are able to claim with certainty, that there is some potential predictability in the staying duration of visits to an area. Therefore, it does make sense to calculate the limits of this predictability.



(a) Beijing (b) Shanghai

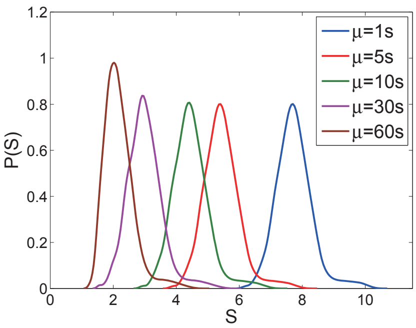
Fig. 2. Average staying duration of one single area in *Beijing* and *Shanghai* datasets.

B. Entropy and Limits of Predictability

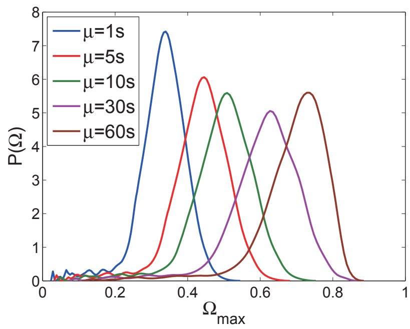
We explore the entropy and limits of predictability in staying duration for all cases. In the slotted cases with τ of 10, 30 and 60 minutes, we calculate the desired value for each time slot of each area. However, in the unslotted cases, we only do the calculation for each area. For every slot time (including unslotted), we further divide it into five cases with μ set to 1, 5, 10, 30 and 60 seconds. Therefore, we do calculation for $4 \times 5 = 20$ cases in total.

We first study the precision μ's influence on entropy S and limit of predictability Ω_{max}. Holding the slot time τ constant, we plot the distributions of S and Ω_{max} with different μ together. The results for the unslotted case are shown in Fig. 3. We notice unidirectional shifts of distributions with the increasing of μ in both datasets. Theoretically, the larger μ is, the less the number of legal staying duration is, the smaller S is, and the higher Ω_{max} is. We can see that the calculation results fit this relationship quite well. We then plot the variation tendencies of Ω_{max} along with μ for different slot time τ in Fig. 4. Clearly, the tendencies are almost the same for different τ.

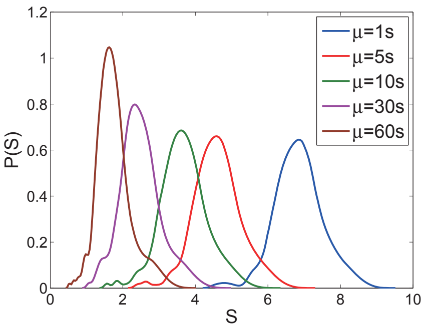
Next, we evaluate the effect of slot time τ on Ω_{max}. With μ fixed, we compare the distributions of Ω_{max} with different τ. In Fig. 5, we can see that a higher predictability can be attained when time slots are used, and Ω_{max} grows with the decrease of slot time τ. This proves that dividing one day into different time slots can help us better predict the staying duration of visits to each area. Here we have to explain that although a higher Ω_{max} is attained for a smaller τ, this does not mean that we should use a slot time as small as possible.



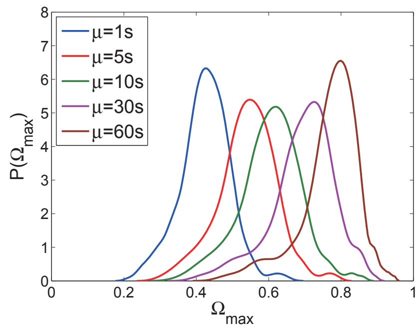
(a) S in Beijing



(b) Ω_{max} in Beijing

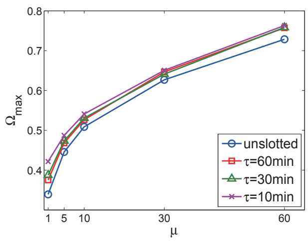


(c) S in Shanghai

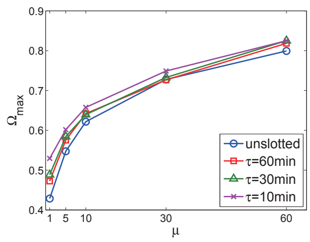


(d) Ω_{max} in Shanghai

Fig. 3. Distributions of Entropy S and limits of predictability Ω_{max} of the unslotted cases with different precision μ in *Beijing* and *Shanghai* datasets.

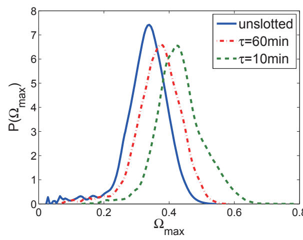


(a) Beijing

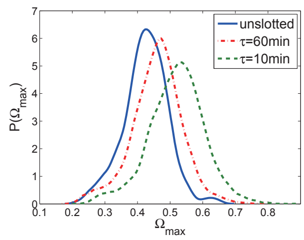


(b) Shanghai

Fig. 4. Variation tendencies of limits of predictability Ω_{max} along with precision μ for different slot time τ in *Beijing* and *Shanghai* datasets.

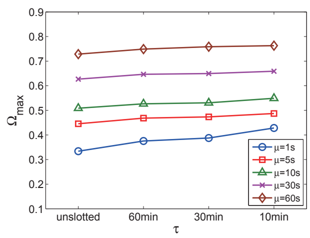


(a) Beijing

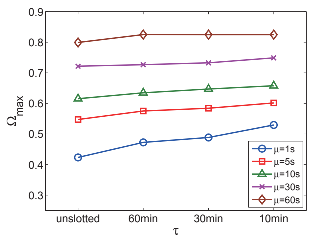


(b) Shanghai

Fig. 5. Distributions of limits of predictability Ω_{max} with $\mu = 1s$ for different slot time τ in *Beijing* and *Shanghai* datasets.



(a) Beijing



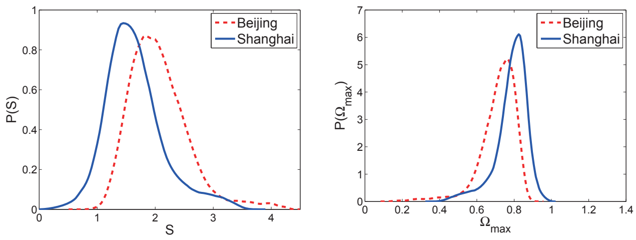
(b) Shanghai

Fig. 6. Variation tendencies of limits of predictability Ω along with slot time τ for different precision μ in *Beijing* and *Shanghai* datasets.

For a τ too small, there may be too few previous visits to the areas which we can use as history for the prediction. And to make matters worse, the slight fluctuation may invalidate the assumption that the staying duration of an area follows the same pattern in each time slot for different days. Therefore, a proper slot time must be carefully selected to maximize the predictability within the allowed range. We leave this question for future work as it is not the focus of this paper.

To find the variation tendencies of Ω_{max} along with τ for different precision μ, we compare the results in Fig. 6. It can be seen that the smaller μ is, the more Ω_{max} can be improved when using time slot and reducing the slot time τ. This is valid because Ω_{max} grows together with μ, which makes the potentiality to further improve Ω_{max} getting smaller. Similarly, we can make choices on whether to use time slot and how much τ should be based on the improvement of Ω_{max} after the precision μ is set as required.

According to the tendencies stated above, we get the small-



(a) S (b) Ω_{max}

Fig. 7. Distributions of Entropy S and limits of predictability Ω_{max} with $\tau = 10min$ and $\mu = 60s$ in *Beijing* and *Shanghai* datasets.

est entropy S and the highest limit of predictability Ω_{max} in the case with $\tau = 10min$ and $\mu = 60s$. In this case, we plot the $P(S)$ of both datasets together in Fig. 7(a) and $P(\Omega_{max})$ in Fig. 7(b) for further analysis and comparison. In Fig. 7(a), $P(S)$ peaks at $S \approx 1.92$ and $S \approx 1.49$ in *Beijing* and *Shanghai* respectively. That is to say, on average, we are able to narrow down the range of possible staying duration of visits to an area to $2^{1.92} \approx 3.78$ and $2^{1.49} \approx 2.81$. In Fig. 7(b), the peak of $P(\Omega_{max})$ is located at 0.763 in *Beijing* and 0.825 in *Shanghai*, which implies that, respectively, 76.3% and 82.5% of the visits to an area last for a duration that may be predicted. Meanwhile, we can also conclude that vehicles in Shanghai have higher potential predictability than ones in Beijing, which accords with the common sense that drivers in Shanghai are better-behaved on the road than those in Beijing.

In the above analysis, we find that lowering the requirement for precision and using small time slot within allowed range can both raise the potential predictability. We get the highest potential predictability of 76.3% in Beijing and 82.5% in Shanghai when we use smallest slot time and lowest requirement for precision among all our cases and conclude that vehicles in Shanghai are more predictable than that in Beijing in general.

VII. Conclusions

In this paper, we use two datasets of vehicular traces collected in Shanghai and Beijing to study the time patterns and limits of predictability in staying duration of vehicles' visits to different areas. We calculate the limits of predictability in many cases to explore the effect of different precision and time slot on predictability. The results showed that a higher predictability can be achieved if the requirement for precision is reduced, and using time slot is an efficient way of prediction.

There are still many problems remained for further solutions. How to select the optimal precision and slot time is a question worth future discussion. Based on these optimal values, a more effective predictive algorithm for staying duration may be proposed to attain the limits of predictability we calculated.

References

[1] A. Stathopoulos and M. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 121–135, Feb. 2003.

[2] Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: comparison of multivariate and univariate approaches," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1857, no. -1, pp. 74–84, Jan. 2003.

[3] G. Dimitrakopoulos and P. Demestichas, "Intelligent Transportation Systems," *IEEE Vehicular Technology Magazine*, vol. 5, no. 1, pp. 77–84, Mar. 2010.

[4] J. Zhao and G. Cao, "VADD: Vehicle-assisted data delivery in vehicular ad hoc networks," in *Proc. 25th IEEE INFOCOM* (Barcelona, Spain), Apr. 23-29, 2006, pp. 1–12.

[5] C. Wu, J. Ho, and D. Lee, "Travel-time prediction with support vector regression," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 5, no. 4, pp. 276–281, Apr. 2004.

[6] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," *IEEE Vehicular Technology Magazine*, vol. 2, no. 2, pp. 12–22, Jun. 2007.

[7] B. Zhang, K. Xing, X. Cheng, L. Huang, and R. Bie, "Traffic clustering and online traffic prediction in vehicle networks: A social influence perspective." in *Proc. 31th IEEE INFOCOM* (Orlando, Florida USA), Mar. 25-30, 2012, pp. 1–9.

[8] V. Namboodiri and L. Gao, "Prediction-based routing for vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 2332–2345, Apr. 2007.

[9] K.A. Ali, M. Lalam, L. Moalic, and O. Baala, "V-MBMM: Vehicular Mask-Based Mobility Model," in *Proc. ICN'10* (Menuires, The Three Valleys, French Alps, France), Apr. 11-16, 2010, pp. 243–248.

[10] H. Huang, Y. Zhu, X. Li, et al, "META: a Mobility Model of Metropolitan Taxis Extracted from GPS Traces," in *Proc. IEEE WCNC'10* (Sydney, Australia), Apr. 18-21, 2010, pp. 1–6.

[11] G. Mohimani, F. Ashtiani, A. Javanmard, and M. Hamdi, "Mobility modeling, spatial traffic distribution, and probability of connectivity for sparse and dense vehicular ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1998–2007, May 2009.

[12] F. Karnadi, Z. Mo, K. Lan, "Rapid Generation of Realistic Mobility Models for VANET," in *Proc. IEEE WCNC'07* (Hong Kong, China), Mar. 11-15, 2007, pp. 2506–2511.

[13] F.J. Martinez, J.C. Cano, C.T. Calafate, and P. Manzoni, "Citymob: A mobility model pattern generator for VANETs," in *Proc. IEEE ICC'08* (Beijing, China), May 19-23, 2008, pp. 370–374.

[14] H. Menouar, M. Lenardi, and F. Filali, "Improving Proactive Routing in VANETs with the MOPR Movement Prediction Framework," in *Proc. 7th ITST* (Sophia Antipolis, France), Jun. 6-8, 2007, pp. 1–6.

[15] S. Wan, J. Tang, and R.S. Wolff, "Reliable Routing for Roadside to Vehicle Communications in Rural Areas," in *Proc. IEEE ICC'08* (Beijing, China), May 19-23, 2008, pp. 3017–3021.

[16] M. Kamilov, I. Karimov, J. Song, H. Kim, and S. Han, "A Real-time Routing System for Vehicular Ad Hoc Networks," in *Proc. STFSSD'09* (Tokyo, Japan), Mar. 17, 2009, pp. 88–92.

[17] C. Song, Z. Qu, N. Blumm, and A.L. Barabsi, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, Feb. 2010.

[18] H. Zhu, M. Li, L. Fu, G. Xue, Y. Zhu, and L. Ni, "Impact of traffic influxes: Revealing exponential inter-contact time in urban VANETs," *IEEE Trans. Parallel and Distributed Systems*, vol. 22, no. 8, pp. 1258–1266, Aug. 2011.

[19] A. Franz, "Voronoi DiagramsCa survey of a fundamental geometric data structur," *ACM Computing Surveys*, vol. 23, no. 3, pp. 345–405, Sept. 1991.

[20] I. Shunsuke, *Information theory for continuous systems* (World Scientific, 1993), p. 2.

[21] A. Brabazon, M. O'Neill, *Natural Computing in Computational Finance* (Springer, Berlin, 2008), vol .100, pp. 197–210.

[22] I. Kontoyiannis, P.H. Algoet, Yu.M. Suhov, A,J. Wyner, "Nonparametric Entropy Estimation for Stationary Processes and Random Fields, with Applications to English Text," *IEEE Trans. Information Theory*, vol. 44, no. 3, pp. 1319-1327, May 1998.