

Traffic-Known Urban Vehicular Route Prediction Based on Partial Mobility Patterns

Zhongwei Li, Guangtao Xue, and Hongzi Zhu
Department of Computer Science & Engineering
Shanghai Jiao Tong University
{lizhongwei, gt_xue, hongzi}@sjtu.edu.cn

Yunhuai Liu
Department of Computer Science & Engineering
Hong Kong University of Science and Technology
yunhuai@cse.ust.hk

Abstract

Travel route analysis and prediction are essential for the success of many applications in Vehicular Ad-hoc Networks (VANETs). Yet it is quite challenging to make accurate route prediction for general vehicles in urban settings due to several practical issues such as very complicated traffic networks, the highly dynamic real-time traffic conditions and their interaction with drivers' route selections. In this paper, we undertake a systematic study on the vehicular route prediction in urban environments where the traffic conditions on complicated road networks keep changing from time to time. Inspired by the observation that a vehicle often has its own route selection flavor when traversing between its sources and destinations, we define a mobility pattern as a consecutive series of road segment selections that exhibit frequent appearance along all the itineraries of the vehicle. We further leverage Variable-order Markov Models (VMMs) to mine mobility patterns from the real taxi GPS trace data collected in Shanghai. In addition, considering the tremendous impact of dynamic traffic conditions to the accuracy of route prediction, we deploy multiple VMMs differentiating different traffic conditions in day-time. Our extensive trace-driven simulation results show that notable patterns can be mined from routes of common vehicles though they usually have no constraints when selecting routes. Given a specific taxi, around 40% next road segments are predictable using our model with a confidence weight of 60%. With multiple VMMs a high route prediction accuracy is achievable from the real traffic trace.

1 Introduction

As the rapid development of wireless technology, more and more studies on vehicular ad hoc network (VANET) are trying to build innovative applications to provide safety and comfort for passengers and drivers, such as fast dissemination of warning messages of traffic accidents, congestion avoidance, sharing media between vehicles and onboard Internet access. To realize these stunning applications, manipulating effective opportunistic wireless communications between different vehicles and vehicles to roadside infrastructure is crucial. For example, two vehicles can communicate with each other in a very short time only when they geographically meet. Therefore, knowing possible locations of vehicles in the future can largely leverage the performance of the whole system. Consequently, it is highly important to accurately predict travel routes of vehicles.

Accurate travel route prediction in urban vehicular environments, however, is very challenging due to the following three reasons. First, the structure of the urban road networks is very complicated. For example, in Shanghai, the largest metropolis in China, there are about 22,413 intersections connecting about 33,290 surface road segments in addition to many tunnels and densely covered viaducts. How to establish the next choice on the path of a vehicle in such intricate settings is not straightforward. Second, a vehicle may be heading for different destinations. Obviously different destinations will lead to different itinerary choices. This adds another dimension of uncertainty to the route prediction problem. Third, traffic conditions are time-varying and will influence the route decision of a vehicle dramatically. Generally speaking, in off-peak hours, drivers would prefer routes of shorter distance but are more likely to choose

routes of low traffic during rush hours.

There has been a large number of previous work in route prediction. Krumm *et al.* [1, 2] proposed to use simple or hidden Markov model (HMM) to predict the short-term driving route of vehicles. Their work mainly focused on the static route selection that took no dynamic traffic conditions into account. Related work in [3] has showed that the driver's route choices are under the influence of real-time traffic conditions. In additional, the vehicle route choice is a very complex process. It is influenced by the structure of the urban road, real-time traffic condition and emergencies on the road. It is quite hard, if not impossible, to obtain the preknowledge about how many immediately preceding road segments are associated with the upcoming itinerary. Simple Markov model is poor in capturing variable order Markov dependencies. Although HMM is capable of handling the vehicles moving patterns, training the HMM model suffers from known learnability hardness [4].

In this paper, we propose a systematic approach to predict travel routes of a vehicle in urban settings. Our approach is made up of three components. First, by analyzing real taxi GPS trace data, we observe that even a specific taxi has evident mobility patterns despite its arbitrary on-demand itineraries. Second, we train separate Variable-order Markov Models (VMMs) [5] for a vehicle according to different traffic conditions (i.e., in rush hours and off-peak hours) in a day, using real GPS trace data. A mobility pattern is extracted if the probability of a road segment sequence presented in a VMM is higher than a threshold. Third, we further exploit Probabilistic Suffix Tree (PST) [6, 7] to predict driving routes based on the retrieved mobility patterns. We conduct extensive trace-driven simulations to verify the validity of our route prediction method and the results show that our methodology can reach high accuracy under different traffic conditions.

The original contributions that we have made in the paper are highlighted as follows:

- We observe that vehicles have obvious mobility patterns in urban settings. After analyzing a large amount of real GPS trace data of about 4,000 taxies, we noticed that, even for a specific taxi, around 40% next road segments on its itinerary are applicable to our extracted mobility patterns with the confidence weight set to 60%.
- We take into consideration the enormous impact of different traffic conditions to the route prediction. Highly dynamic traffic conditions can greatly affect route choice of a vehicle. We deal with dynamic traffic condition by training different VMMs and hence can obtain high prediction accuracy under different traffic conditions.
- We leverage VMMs to characterize the inner correlation between road segments in the itineraries of a vehicle and extract those with strong correlation as mobility patterns.
- We conduct extensive trace-driven simulations to verify the efficacy of our route prediction approach. Experiment results show our approach can reach high route prediction accuracy in the urban environments.

The rest of the paper is organized as follows. In Sec. 2, we give a brief summary of previous related work. In Sec. 3, we describe our map-matching method for processing the raw GPS traffic data. In Sec. 4, we introduce the extraction of mobility patterns using VMMs, and describe the route prediction algorithm in Sec. 5. Following that we discuss the experimental results in Sec. 6. Finally, we conclude our study and shin section 7.

2 Related Works

Previous studies [8,9] have emphasized that vehicles motion characteristics have obvious impact on VANET performance. To utilize the benefit of vehicles motion characteristics, many researchers on VANET have paid their attentions on modeling vehicles underlying characteristics [10–12], and vehicles moving regularity is one important aspect of these characteristics.

Regularity of vehicles behavior lets us be able to predict moving vehicles future destinations with a certain degree of probability. Karbassi *et al.* [13] attempts to predict the vehicles traveling paths between given origin and destination. Froehlich *et al.* [14] tries to find the drivers regular routes and predict which roads the drivers are currently driving on. Studies in [1, 2] are most closely related to our work. These two studies are both focusing on the short-term upcoming roads prediction without concerning the traveling route starting and destination. However, they just simply evaluate the temporary and velocity impact on driving route prediction, leading to the loss of improvement the prediction accuracy by considering the different traffic conditions during one day.

Different prediction models have been used for generating VMP or predicting the vehicles future destinations. Studies in [12, 15] use Kalman filtering method to predict users potential destinations [12] or forecast traffic volume [15]. Gehrke *et al.* [16] uses AQ21, which is a natural induction system that learns rules and applies rules to make traffic prediction. In studies [1, 2], simple Markov model and HMM are used to predict the drivers' future roads respectively. In our work, we use VMM to model the vehicles moving patterns, and apply PST to make driving route prediction. Since vehicle moving patterns are of variable



Figure 1. A taxi with a commercial GPS device installed, the highlight area in the inset shows such a device

length, our methodology is more adaptive and convenient to be implemented.

3 Real Traffic Traces Processing

We collected real Taxi GPS data from the ShanghaiGrid project [17]. By this project commercial GPS receivers are installed to certain public vehicles (around 4,000 taxis and 1,000 buses). In general, taxis have much more mobility and higher dynamics than common vehicles due to different demands of passengers. The study of route prediction on taxis enlightens route prediction on other regular vehicles which generally have more obvious mobility patterns.

In Fig. 1, a taxi with a GPS receiver is shown. A vehicle actively reports its current status information back to a data center through a wireless cell-phone data channel (i.e., GPRS). Due to the GPRS communication cost for transmitting, drivers prefer to choose relatively large intervals. The typical value is one minute. The information we can obtain directly from GPS reports is very limited: a vehicle's location coordinates, timestamp, and the optional speed and heading.

To study mobility patterns, we first need to tell which road a taxi is actually running on. Thus, we need to recover each sample back on track. There are two major causes that make this procedure non-trivial. First, the reported GPS coordinates are very erroneous in the city setting with dense high buildings and viaducts. For example, the error of reported locations can be as large as 100 meters. Second, the relatively long report intervals add much difficulty to determine the precise route between two single reports. A taxi can actually runs out several road segments in such a long period of time when the traffic conditions are good.

To tackle these problems, we implement a map-matching algorithm which takes into account not only the location deviation between GPS reports from a taxi and possible road segments but also the mobile context of this taxi. To recover a report, the algorithm prefers those roads with minimum projection distance from the report and minimum angle de-

```

Mobility Pattern Generation Algorithm
Phase 1 --- Initialization
  Let  $\bar{T}$  consist of the single root node and  $\bar{S} = \Phi$ 
Phase 2 --- Building PST with two orders
  1. Divide the traces and direct adjacent roads into
     two classes according vehicle moving direction
  2. For each class, pick  $\forall r_n$  from direct adjacent
     roads
  3. If  $\bar{P}(r_n|r_c) > \sigma$  and  $\bar{N}_{rc} > N_{min}$ 
     then add to  $\bar{T}$  the node  $r_n r_c$ ,
     add  $s = r_n r_c$  to
  5. End For
Phase 3 --- Building PST with more layers
  1. While  $\bar{S} \neq \Phi$ , pick any  $s_c \in \bar{S}$  do
  2. Remove  $s_c = \{r_n r_c r'_1 r'_2 \dots r'_k\}$  from  $\bar{S}$ 
  3. For road  $r'_{k+1}$  of  $r'_k$  adjacent roads and  $r'_{k+1}$ 
     not in  $r_c r_1 r_2 \dots r_k$ 
  4.   If  $\bar{N}_{sc} > \bar{N}_{min}$  and  $>$ 
  5.     Then add to  $\bar{T}$  the node  $r_n r_c r'_1 r'_2 \dots r'_k r'_{k+1}$ 
  6.     If  $k+1 < K$ 
  7.       Then add  $s = \{r_n r_c r'_1 r'_2 \dots r'_k r'_{k+1}\}$  to
  8.     End For
  9. End While

```

Figure 2. Mobility pattern generating algorithm

viation between the heading direction of the taxi and the road. This simple yet effective strategy works well in most situations and can gain very high accuracy compared with real itineraries. In more complicated cases where the geographical distance between these two consecutive records could exceed thousands of meters, we need to consider the mobile context of the taxi. The algorithm examines several previous and successive reports to determine the most possible road segment where the report issued. Our on-road experiment results show that our map-matching algorithm can reach about 98% accuracy.

4 Extracting Vehicular Mobility Patterns

After processing the raw taxi GPS trace data, the geographical locations of taxis are all adjusted to the corresponding roads. Let $R = \{R_1, R_2, \dots, R_n\}$ be the set of all road segments and R_i as the i^{th} road segment. We present a m -segment itinerary as a road segment sequence containing m consecutive road segments, denoted as $S^m = r_1 r_2 \dots r_m$ where $r_i \in R, m \geq 1$.

Definition 1: We define a k -segment Vehicular Mobility Pattern ($k \geq 2$) as a k -segment itinerary S^k where, given all $S^j, 1 \leq j < k$, the probability of a taxi will select r_{j+1} as the next route choice is no less than a predefined threshold σ .

4.1 Training variable-order Markov model

Through observation the vehicles moving trajectories, we find that the vehicles motion has the feature, termed short memory, which is common to many natural sources. The short memory feature indicates that for a certain road segment sequence the probability distribution of the next road segment given the preceding subsequence can be quite accurately approximated by observing no more than the last K road segments in that subsequence.

In our route prediction approach, we deploy PST which in essence is an implementation of K -bounded VMM (K is the memory length of the model). In a K -bounded VMM, a state can be determined by knowing at most K previous states in the Markov chains. To mine mobility patterns of a vehicle, we construct a PST for each individual road as follows:

Step 1. Vehicle traces parsing. For each road the vehicle passed, only the K -bounded preceding roads and the next adjacent road need to be considered for mining VMP. This step extracts the relative road segment sequences for each road the vehicle has passed.

Step 2. Building the PST. Based on the passed road segment sequences, the PST is constructed for each road r_c . Let \bar{T} denote the tree, $\bar{S} = \{s_1, s_2, \dots, s_n\}$ denote the set of subsequences that need to check, and $\bar{N}s$ denote the number of distinct subsequences which include subsequence s . We also define two thresholds for VMP generation. N_{min} denotes the minimum allowed value for $\bar{N}s$, and σ denotes the required empirical probability \bar{P} [7] for VMP generation. First, we create a PST consisting of a single root node. Second, for each road r_n in the direct adjacent roads of r_c , if empirical probability is bigger than σ , the subsequence $r_n r_c$ is treated as a VMP and the node $r_n r_c$ is added to \bar{T} . Also, the subsequence $s = r_n r_c$ is added into for higher-order VMP mining. To enable directional VMP mining, the vehicle moving direction information is used to partition the direct adjacent roads and trace subsequences into two classes, and empirical probability calculating is executed independently in each class. Third, for each $s = \{r_n r_c r'_1 r'_2 \dots r'_k\}$ of \bar{S} , if the empirical probability is bigger than $\bar{P}(r_n | r'_c r'_1 r'_2 \dots r'_k)$, then the subsequence $s = \{r_n r_c r'_1 r'_2 \dots r'_k r'_{k+1}\}$ is also a VMP, and the node $r_n r_c r'_1 r'_2 \dots r'_k r'_{k+1}$ is added to \bar{T} . If $k + 1$ is smaller than K , then the subsequence $s = \{r_n r_c r'_1 r'_2 \dots r'_k r'_{k+1}\}$ is continued to be added to. Here, notice that r'_{k+1} is not the next road of r'_k , but the preceding road of r'_k . During this process, for each possible subsequence s , the value of $\bar{N}s$ should be bigger than N_{min} . Since we make prediction only when vehicles on the roads with VMP, the prediction probabilities smoothing step is unnecessary in our algorithm. Fig. 2 gives the pseudo code of our VMP generation algorithm.

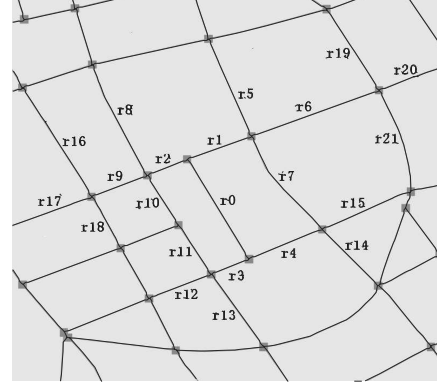


Figure 3. One taxi VMP distribution in Shanghai

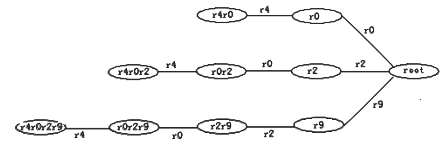


Figure 4. An instance of PST

Fig. 3 shows an instance of road map around road r_0 . Suppose the vehicle historical road segment sequences around r_0 are: $r_{16}r_9r_2r_0r_4r_{15}r_{21}$, $r_{17}r_9r_2r_0r_4r_{15}$, $r_{18}r_9r_2r_0r_4r_{15}$, $r_{18}r_9r_2r_0r_4r_{15}$, $r_8r_2r_0r_3r_{13}$ and $r_5r_1r_0r_3r_{13}$. The valid subsequences for r_0 are $r_9r_2r_0r_4$, $r_9r_2r_0r_4$, $r_9r_2r_0r_4$, $r_9r_2r_0r_4$, $r_8r_2r_0r_3$ and $r_5r_1r_0r_3$ where $K = 2$. Let $\sigma = 0.6$ and $N_{min} = 2$. We get the following equations that satisfy the VMP generation conditions.

$$\begin{aligned}\overline{P}(r_4|r_0) &= 4/6 \approx 0.67 \\ \overline{P}(r_4|r_0r_2) &= 4/5 = 0.8 \\ \overline{P}(r_4|r_0r_2r_9) &= 4/4 = 1.0\end{aligned}$$

Therefore, the road segments of the sequences $r_9r_2r_0r_4$, $r_2r_0r_4$ and r_0r_4 are regarded being high correlated and treated as the vehicle VMPs on road r_0 . Fig.4 shows the corresponding PST on road r_0 .

4.2 Impact of traffic conditions

Vehicle moving behavior is part of human social behavior. To our direct impression, it would be influenced by people daily manners seriously, leading to dynamic traffic conditions in one day. For example, people transport to job in the morning and to family at dusk during the rush hour in the weekday, easily leading to the traffic congestion at these specific time periods.

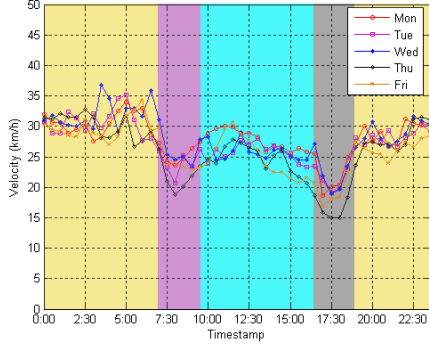


Figure 5. An example of dynamic traffic condition changes

To investigate the traffic condition changes during one day, the vehicle moving speed is chosen as the measurement criterion, due to the fact that it can be as the indicator of the dynamic traffic conditions. We analyze the vehicle moving speed at a large amount of major roads in the urban area of Shanghai. Fig. 5 shows the change of vehicle moving speed during one day at Zhojiabang Rd, which is the major road in the downtown area of Shanghai. From Fig. 5, we can know that the vehicle moving speed decrease obviously at the time periods around 7:00~9:30 and 16:30~19:00, which are corresponding to the rush hour.

Previous study [18] points out that the human mobility has a great degree of temporary regularity and the work in [3] shows the real-time traffic condition would influence the drivers route choices. To gain a detailed knowledge of the traffic condition impact on VMP, we divide one day into several time periods according to the change of vehicles moving speed. We define two basic rules for the time periods partition: traffic condition should be apparently different between adjacent time periods, and the number of time periods should be as small as possible. According to these two rules, we divide one day into four time periods, as Fig.5 shows.

Accordingly, we divide the vehicles historical traces into several parts, and regenerate VMPs with these partial traces for special time periods. Comparing the VMP in different time periods, we find that vehicles may change their driving routes on some roads in different time periods, as Fig. 6 shows. In Fig. 6, during the rush hour, the vehicle moves along the Nanqichang Rd, on which the traffic flow volume is low. But during off-peak hour, the vehicle would prefer turning from the Nanqichang Rd to Xilinbei Rd, around which the road conditions are better. The diversification among VMPs in different time periods means that, given a special pair of origin and destination, vehicles maybe prefer to choose different driving routes under different traffic conditions. This phenomenon implies that considering the



Figure 6. VMPs under different traffic conditions

traffic conditions can help to improve the accuracy of driving route prediction.

5 Predicting Routes with Mobility Patterns

After training the Markov model using the vehicle historical traces, the model can be used to predict the vehicle next road based on the preceding roads the vehicle just passed. As the algorithm shown in the last section, for a vehicle, the roads on which the vehicle has mobility patterns would have their own PST.

When the vehicle moves to road r_c at time t , and the roads it just passed are $r_1 r_2 \dots r_k (k < K)$. To predict its next road, it checks whether a PST was constructed for r_c for the corresponding time period. If no such a PST was found, the vehicle would give up predicting its next road. Otherwise, it search the PST to find the node with the longest suffix matching with $r_c r_1 r_2 \dots r_k$ and ending just with r_n , where r_n is one of the unidirectional adjacent road of r_c . For example, in Fig. 4, if the road segment sequence is $r_2 r_0 r_4$, the pattern $r_2 r_0 r_4$ would be chose, and if the road segment sequence is $r_9 r_2 r_0$, the selected pattern would be $r_9 r_2 r_0 r_4$. If there is such a node, then the vehicle predicts that r_n is the next road that the vehicle would move to. If no such a node is found, it means that selection of the vehicle next road has large randomness, and the vehicle cannot apply VMP on road r_c in its moving direction.

Obviously, the prediction method based on PST is convenient to decide which pattern to be applied when the vehicle has multiple patterns on the same road. It releases us from trivial compassion among all the potential patterns, and always returns us the pattern with the highest probability.

6 Performance Evaluation and Results

In this section, we provide the performance evaluation of our proposals. The experiments are based on the taxis trace data from September, 2006 to April, 2007. The traces in the

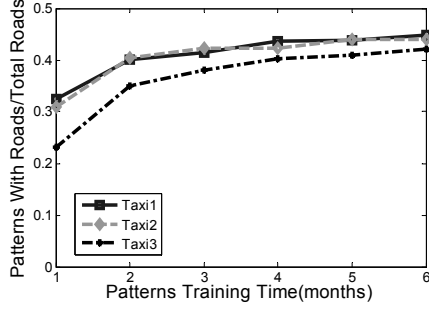


Figure 7. Pattern utilization

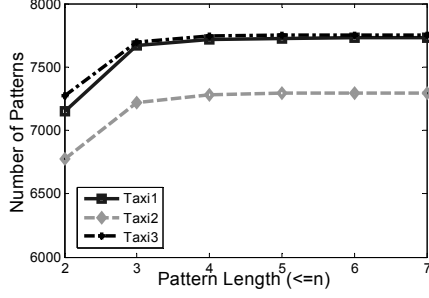


Figure 8. Distribution of Pattern length

first six months are used for the VMM model training, and the last month traces are used for VMP prediction accuracy evaluation.

Our evaluation is performed under Linux system with the GTK graphics library support. The pattern generating threshold σ is set to the value bigger than 0.5 to make sure that only one next road selection is made when applying VMP for prediction. The memory length K of VMM is set to 3.

6.1 Metric definition

The performance of driving route prediction methodologies depend on the ratio of successful predictions (prediction accuracy) and the frequency of applying the VMP for prediction (pattern utilization). These two metrics are defined as follows:

- **Prediction accuracy:** When a taxi has VMP on some roads, it applies the VMP to predict the next driving route. The prediction accuracy defines the ratio of the number of successful predictions to the total number of predictions.
- **Pattern utilization:** Given a k -segment itinerary of a taxi, we define *Pattern Utilization* as the ratio of the total number of mobility patterns used during the taxi's movement along the itinerary to the number of all road segments.

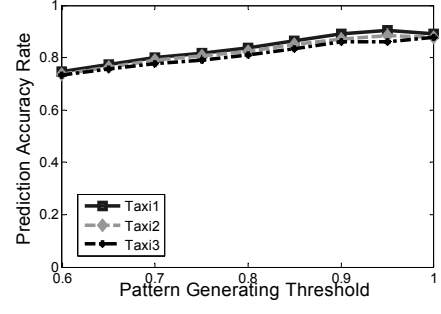


Figure 9. Prediction accuracy

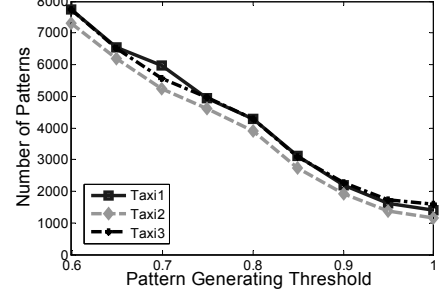


Figure 10. Pattern threshold selection

6.2 Impact of model training time

To well measure *VMP*, one aspect that needs to be considered is how often *VMP* can be used. Higher pattern utilization (PU) means greater proportion of roads are enable to apply *VMP* for predicting the driving route according to the past routes.

Fig. 7 shows the PU of selected taxis based on the taxis traces. Obviously, as the training time increases, more *VMP* can be discovered and PU goes up. As illustrated in Fig.7, with the time period of training time increases from one month to six months, PU of taxi1 grows up to 44.88% (with pattern generating threshold 0.6).

6.3 Impact of memory length K

Vehicle upcoming destinations are seriously influenced by the roads the vehicle just passed. To study the memory length of vehicle motion and determine the most suitable value of K for VMP generation, we analyze the length distribution of VMP.

Fig. 8 plots the patterns' length distribution, which indicates that a majority of *VMP* is of length 2 to 3. As the value of pattern length gets larger than 5, the number of increased patterns almost approaches to zero, which will be ignored in order to reduce the computing complexity. Hence we set the value of K to be 3.

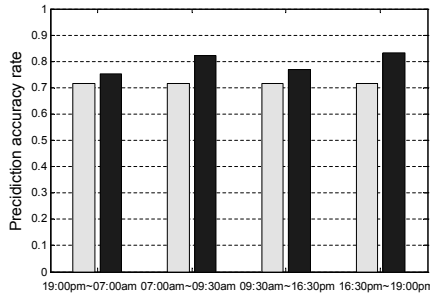


Figure 11. Prediction accuracy comparing between VMP with consideration of traffic conditions (blue bar) and VMP without consideration of traffic-conditions (yellow bar)

6.4 Pattern generating threshold

The pattern generating threshold is an important parameter in our methodology. It can directly influence the prediction accuracy and PU. Selecting the pattern generating threshold carefully can help to improve our model's performance.

Fig. 9 shows various prediction accuracies with different pattern generating thresholds. Generally, as the pattern generating threshold increases, the prediction accuracy increases accordingly. However, for taxi 1 and taxi 3, the prediction accuracy with threshold 1.0 is lower than that with threshold 0.9. The reason is that the cardinal number for generating the patterns with the threshold 1.0 is so small that even one prediction error may decrease the prediction accuracy a lot. The patterns generating threshold decides the number of patterns generated.

In Fig. 10, the number of patterns declines rapidly as the pattern generating threshold increases. While larger threshold means higher prediction accuracy, it also means fewer *VMP* can be applied for prediction. Clearly, there is a trade-off in the selection of threshold. However, as the amount of training data increase, our prediction scheme will generate more *VMP*. Thus we can increase the patterns generating threshold gradually to get more accurate prediction.

6.5 Traffic conditions

Fig.11. Prediction accuracy comparing between VMP with consideration of traffic conditions (blue bar) and VMP without consideration of traffic-conditions (yellow bar)

To validate the impact of traffic condition on VMP, we separate the vehicles traces into four parts according to the time periods shown in Fig.5, and generating VMPs individually from each partial trace data. Fig.11 shows the prediction accuracy contrast between VMP with consideration of

traffic conditions and without consideration of traffic conditions using the traces from September of 2006 to March of 2007. The VMP generating threshold σ is set to 0.6 in our experiment. From the results, we conclude that VMP with consideration of traffic conditions are more suitable to forecast the vehicle driving routes.

7 Conclusions and Future Work

In this paper, we study the mobility patterns of general vehicles in urban settings. We use the real traces collected from Shanghai taxis and propose a VMM to formalize the extraction and prediction of vehicle short-term route prediction. Through the experiments we find that though general vehicles typically have no constraints on route selection, which is inherently different from public transport vehicles (e.g., bus), their actually route selection is also well behaved that a great deal of patterns can be mined. More over, the patterns are highly connected with the real-time traffic conditions. As long as the source, destination and the real-time traffic conditions are available, the prediction accuracy of the next road segment can be greatly increased. More over, when more data are available for model training, the applicability of the model is also notably increased. With six months data, the pattern utilization is increased to 44% compared with that of 22% which is based on one month data.

Our future work will be carried out along following directions. First, we will see how to leverage these well-behaved patterns to facilitate communications in VANETs. When the routes of vehicles are available, the vehicle's inter-contact is also largely available combining the real-time traffic condition. Such knowledge will bring new benefits for communication protocol design in VANETs. Second, the current approach is centralized, which present poor scalability when more vehicles are involved in. We are thinking of distributed algorithm which is portable to every vehicle for their prediction. Third, we are thinking of more advanced transportation management with the mobility patterns when they are available.

References

- [1] Simmons, R. Browning, B. Yilu Zhang, Sadekar, V. "Learning to Predict Driver Route and Destination Intent", Intelligent Transportation Systems Conference, 2006, pp127-132
- [2] J. Krumm. A Markov Model for Driver Turn Prediction. SAE World Congress, 2008
- [3] Yisong Zheng, Qian Wang, Wangling and Mo Kuang. Research into the driver's route choice under ex-

- isting real-time traffic information. *Industrial Engineering and Engineering Management*, 2008. IEEM 2008. IEEE International Conference on IEEM 2008, pp.1515-1518.
- [4] Gillman. D and Sipser. M, Inference and minimization of hidden Markov chains. In *Proceedings of the 7th Annual Workshop on Computational Learning Theory*, 1994, pp.147-158
 - [5] R Begleiter, R El-Yaniv, G Yona. On Prediction Using Variable Order Markov Models. *Journal of Artificial Intelligence Research*, 2004, pp.385-421
 - [6] Dana Ron, Yoram Singer and Naftali Tishby. The Power of Amnesia: Learning Probabilistic Automata with Variable Memory Length. *Machine Learning*, 1996, pp.117-149
 - [7] Gill Bejerano and Golan Yona. Variations on probabilistic suffix trees: statistical modeling and prediction of protein families. *Bioinformatics*, 2001, pp.23-43
 - [8] G. Marfia, G. Pau, E. De Sena, E. Giordano, M. Gerla. Evaluating Vehicle Network Strategies for Downtown Portland: opportunistic infrastructure and the importance of realistic mobility models. *The First International Workshop on Mobile Opportunistic Networking ACM/SIGMOBILE MobiOpp*, 2007, pp.47-51
 - [9] G. Marfia, G. Pau, E. De Sena, E. Giordano, M. Gerla. VANET: On Mobility Scenarios and Urban Infrastructure. A Case Study. *Mobile Networking for Vehicular Environments*, 2007, pp.31-36
 - [10] ETH Zurich, Switzerland. A Mobility Model Based on WLAN Traces and its Validation. *INFOCOM 2005.24th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2005, pp.664-674
 - [11] Jungkeun Yoon, Brian D. Noble, Mingyan Liu, Minkyong Kim. Building realistic mobility models from coarse-grained traces. *Proceeding of the 4th international conference on Mobile systems, applications and services*, 2006, pp177-190
 - [12] Kim, M. Kotz, D. Kim, S. Extracting a mobility model from real user traces. *INFOCOM 2006, 25th IEEE International Conference on Computer Communications*, 2006, pp1-13.
 - [13] Karbassi, A. Barth, M. Vehicle route prediction and time of arrival estimation techniques for improved transportation system management. *Intelligent Vehicles Symposium*, 2003, pp.511-516
 - [14] Jon Froehlich, John Krumm, Route Prediction from Trip Observations. *SAE World Congress & Exhibition*, 2008
 - [15] Okutani, I. and Y. Stephanedes. Dynamic Prediction of Traffic Volume through Kalman Filtering Theory, *Transportation Research Record*, 1984, pp.1-11
 - [16] Jan D. Gehrke and Janus Wojtusiak, Traffic Prediction for Agent Route Planning, *Lecture Notes in Computer Science*, 2008, pp.692-701
 - [17] Minglu Li , Min-You Wu, Ying Li, Jian Cao, Jinpeng Huang, Qianni Deng, Xinhua Lin, Changjun Jiang, Weiqin Tong, Yadong Gui, Aoying Zhou, Xinhong Wu, Shui Jiang. *ShanghaiGrid: an Information Service Grid, Concurrency and Computation: Practice and Experience*, Vol.18, pp.111-135
 - [18] M. C. Gonzale, C. A. Hidalgo, and A-L. Barabasi. Understanding Individual Human Mobility Patterns, *Nature*, 2008, pp.779-782