

Study Of Statistical Models For Route Prediction Algorithms In VANET

Uma Nagaraj*,Nivedita N.Kadam
Dept. of Computer Engg. M.A.E , Alandi (D), Pune, India.
* Umanagaraj67@gmail.com

Abstract

Vehicle-to-vehicle communication is a concept greatly studied during the past years. Vehicles equipped with devices capable of short-range wireless connectivity can form a particular mobile ad-hoc network, called a "Vehicular Ad-hoc NETwork" (or VANET). The users of a VANET, drivers or passengers, can be provided with useful information and with a wide range of interesting services. Route prediction is the missing piece in several proposed ideas for intelligent vehicles. In this paper, we are studying the algorithms that predict a vehicle's entire route as it is driven. Such predictions are useful for giving the driver warnings about upcoming traffic hazards or information about upcoming points of interest, including advertising. This paper describes the route Prediction algorithms using Markov Model, Hidden Markov Model (HMM), Variable order Markov model (VMM).

Keywords- VANET, MANET, ITs, GPS, HMM, VMM, PST.

1. Introduction

The increasing demand of wireless communication and the needs of new wireless devices have tend to research on self organizing, self healing networks without the interference of centralized or pre-established infrastructure/authority. The networks with the absence of any centralized or pre-established infrastructure are called Ad hoc networks. Ad hoc Networks are collection of self-governing mobile nodes.

Vehicular Ad hoc Networks (VANET) is the subclass of Mobile Ad Hoc Networks (MANETs). VANET is one of the influencing areas for the improvement of Intelligent Transportation System (ITS) in order to provide safety and comfort to the road users. VANET assists vehicle drivers to communicate and to coordinate among themselves in order to avoid any critical situation through Vehicle to Vehicle communication e.g. road side accidents, traffic jams, speed control, free passage of emergency vehicles and unseen obstacles etc. Besides safety applications VANET also provide comfort applications to the road users. For example, weather information, mobile e-commerce, internet access and other multimedia applications. The most well known applications include, "Advance Driver Assistance Systems (ADASE2) (Heddebaut 2005), Crash Avoidance Matrices Partnership (CAMP) (Shulman 2005), CARTALK2000 (Andreone 2005) and Fleet Net" (Kruger 2005) that were developed under collaboration of various governments and major car manufacturers. Figure 1 shows the overall working structure of VANET.

2. Need of Route Prediction in VANET

Accurate travel route prediction in urban vehicular environments, however, is very challenging due to the following three reasons (Guangtao 2009).

1. The structure of the urban road networks is very complicated. For example, in a city environment there are about 1000 intersections connecting about 25000 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.
2. 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.

3. 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. 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, to obtain the pre knowledge about how many immediately preceding road segments are associated with the upcoming route.

3. Statistical Route Prediction Models

Here in this section we are going to describe the three statistical models for predicting the routes for vehicle driver .For prediction of a future route for vehicle driver we have to attach a GPS to each vehicle and we have to follow the route traces of a vehicle then we have to convert this collected GPS data into the map matching algorithm. Because many times in a urban environment due to the large buildings and road segments covered under viaducts leads to produce the GPS data erroneous, which produces the uncertain route prediction.

3.1 Markov Model

Markov Model predicts only vehicle's near term future route is based on its near term past route (SAE 2008). A Markov model is a graphical statistical model that captures a sequential model of behaviour. It is a tuple < S,A,T >, where S is a (finite) set of states, A is a (finite) set of actions, and T is the transition function $T : S\times A\times S\rightarrow R$, where $T(s_i,a, s_j) = p(s_i|s_j,a)$, which is the probability of transitioning to state s_i given that the system is in state s_j and action a is executed. Given a Markov model and an initial state distribution π, one can predict the state distribution that results from carrying out a sequence of actions $<a_1, a_2, ..., a_n>$. If $p^t(s_i)$ is the probability of being in state si at time t (where $\pi(.) = \pi^0(.)$ is the initial state distribution), then $p^{t+1}(s_i) = \Sigma s_j \epsilon S\ p^t(s)T(s_i,{}^{at+1}, s_j)$. Note that while the exact state of the system is uncertain when doing prediction, it is assumed that the state is known for certain after the actions are actually executed in the world.

The Markov model can be used to predict beyond just the next road segment. We can clearly build $P[X(1)|X(0)]$, where x(0) is the current road segment and x(1) is next road segment, which is the distribution over the road segments after the next one, given the current one. We can also user higher order models to make these farther out predictions, e.g. $P[X(2)|X(-1), X(0)]$. In general, we can build an n^{th} order Markov model ($n \geq 1$) to predict the m^{th} next encountered segment ($m \geq 1$). The general n^{th} order model is denoted as,

$$P_n[X(m)]=P[X(m)\,|\,X(-n+1),X(-n+2),\ldots,X(0)] \qquad (1)$$

Our prediction algorithm is based on observations of where drivers drive measured from GPS receivers. We do not attempt to predict where on the road segment a vehicle will be, nor do we attempt to predict when it will arrive at a road segment. Our goal is predict the chain of road segments that a vehicle will next encounter. In this way, our predictions show the vehicle's future path, elevation, and turns.

It observes at how prediction accuracy changes as it increase n, the number of segments look at into the past (better). It also looks at how prediction accuracy changes as m increases, the number of segments it predicts into the future (worse). The Markov model does not explicitly constrain a vehicle to adhere to the connectedness of the road network. A trained model could conceivably predict that a driver will jump over several road segments. However, since the model is trained from real data, where such jumps do not occur, the Markov model implicitly prevents such nonsense predictions. One advantage of probabilistic predictions is that the algorithm has a measure of its own uncertainty that can be usefully reported to in-vehicle applications. For instance, automatically engaging a turn signal might depend on near 100% prediction certainty, while presenting a point of interest would not require the same level of confidence.

3.2 Hidden Markov Model

A Hidden Markov model (HMM) (Simmons) is a Markov model with hidden (unobservable) state. An HMM is a five tuple < S,A,O,T,Z,p >, where S, A, and T are the same as with the Markov model and p is the initial state distribution. In addition, O is a (finite) set of observations and Z is the observation function $Z : O\times S\times A\rightarrow R$, where Z(o, s, a) =p(o| s, a), which is the probability of receiving observation o given that the system ends up in state s after executing action a. For many problems, Z is the same for all values of a, (i.e., $Z(o, s_i, a_j)=Z(o, s_i, a_k)$). In what follows, it will use Z(o, s) as shorthand for Z(o, s, a), when Z is the same for all values of a.

As in a Markov model, the exact state of the system is uncertain when predicting the effects of actions. Unlike a Markov model, however, the state may remain uncertain even after executing the actions. For a HMM, both transitions and observations are used to help infer the next state distribution. The same equation as given above for Markov models is used to predict the state distribution at time t +1 given the state distribution at time t and the action a^{t+1}. In addition, o^{t+1}, the observation at time t +1 is used to further constrain the state distribution:

$$P^{t+1}(s)\leftarrow p^{t+1}(s)Z(o^{t+1}, s, a^{t+1})/ p(o^{t+1}) \tag{2}$$

Where $p(o^{t+1})$ is a normalizing factor and is given by,

$$p(o^{t+1}) = \sum_{s_i \varepsilon S} p^{t+1}(s_i)z(o^{t+1},s_i,a^{t+1}) \tag{3}$$

3.3 Variable-order Markov Model

VOM models arose as a solution to capture longer regularities while avoiding the size explosion caused by increasing the order of the model. In contrast to the Markov chain models, where each random variable in a sequence with a Markov property depends on a fixed number of random variables, in VOM models this number of conditioning random variables may vary based on the specific observed realization, known as context. These models consider that in realistic settings, there are certain realizations of states (represented by contexts) in which some past states are independent from the future states leading to a great reduction in the number of model parameters.

Algorithm for learning VMM over a finite alphabet ∑. Such algorithms attempt to learn probabilistic finite state automata, which can model sequential data of considerable complexity. In contrast to N-gram Markov models, which attempt to estimate conditional distributions of the form P(σ|s), with S Є ∑N and σ Є ∑, VMM algorithms learn such conditional distributions where context lengths |s| vary in response to the available statistics in the training data. Thus, VMMs provide the means for capturing both large and small order Markov dependencies based on the observed data. Although in general less expressive than HMMs, VMM algorithms have been used to solve many applications with notable success. The simpler nature of VMM methods also makes them amenable for analysis, and some VMM algorithms that we discuss below enjoy tight theoretical performance guarantees, which in general are not possible in learning using HMMs.

For route prediction VMM extends to probabilistic suffix tree (PST) algorithm (Ron et al., 1996) attempts to construct the single “best" D-bounded VMM according to the training sequence. It is assumed that an upper bound D on the Markov order of the “true source" is known to the learner.

A PST over ∑ is a non empty rooted tree, where the degree of each node varies between zero (for leaves) and │∑│each edge in the tree is associated with a unique symbol in ∑. These edge labels define a unique sequence s for each path from a node v to the root. The sequence s labels the node v. Any such PST tree induces a “suffix set" S consisting of the labels of all the nodes. The goal of the PST learning algorithm is to identify a good suffix set S for a PST tree and to assign a probability distribution P(σ|S) over ∑, for each s Є S.

4. Algorithms for Route Prediction

In this section we have described the detailed algorithms for route prediction using statistical models from previous section.

4.1 Route Prediction Algorithm using Markov Model

As described in previous section Markov Model(SAE 2008) is able to predict the next route segment but it has to know the past road segment, which is dependent on its previous road segment.

Steps for route prediction are as follows:

INPUT: Trained vehicle road segments

OUTPUT: Predicts the future road segment

1. Collect the experimental GPS data from the vehicles route segments. (This data is a set of time-stamped latitude and longitude pairs).
2. Convert this GPS vehicle data with road segment with the help of map matching algorithm (Shanugan 1998).
3. This algorithm matches each GPS point to road segment, taking into account which roads are nearby as well as constraints imposed by the connectivity and speed limits of the road network.
4. Process the road segments to eliminate any adjacent repeated road segments. By considering each trip represented by a continuous road segment.

4.2 Route Prediction Algorithm Using Hidden Markov Model

Hidden Markov Model observes the hidden states. It learned from trips. It is described by Ried Simmons, that a trip is a sequence of data points, where each data point contains the link and time stamp indicating when the vehicle was on the link .The link associated with the last data point is considered as a goal of the trip. For a trip from source to destination route segments are divided into number of links. Links are nothing but the road segments.

Steps for route prediction are as follows:

INPUT : Trained vehicle road segments.

OUTPUT: Predicts the future road segments.

1. Collect the experimental GPS data from the vehicles route segments. (This data is a set of time-stamped latitude and longitude pairs).
2. Convert this GPS vehicle data with road segment with the help of map matching algorithm.
3. Go through the trip sequences, whenever there is transition from one link l_j to another l_i to access element l_j in the hash table increment the number of transitions to $l_{i,}$ under the appropriate condition.
4. If goal link (last link)is g then, find the entry < l_i, g, m> exists if found increment m. where m=time, g=goal, l=link

   Otherwise, add a new element to < l_i, g, 1> for the link 1.
5. This process is repeated from g to the END_OF_TRIP link.
6. The statistics for the time spent on l_j are updated, where the time spent on the l_j is determined as the time stamp the first trip data point that is on l_j until the time stamp of the first trip data point that is on $l_{i.}$

It yields accurate prediction is that the training data is reliable.

4.3 Route Prediction Algorithm Using Variable order Markov Model(Guangtao 2009)

INPUT : Trained vehicle road segments.

OUTPUT: Predicts the future road segments.

1. Collect the experimental GPS data from the vehicles route segments. (This data is a set of time-stamped latitude and longitude pairs).
2. Convert this GPS vehicle data with road segment with the help of map matching algorithm.
3. Trace the raw GPS trace data, geographical locations of vehicle are all adjusted to corresponding roads. Let $R=\{R_1,R_2,\ldots R_n\}$ be the set of all road segments and R_i is the i^{th} road segment and m containing the consecutive road segments, as $S^m=r_1r_2,\ldots.r_m$ where $r_i \in R$, m>1.
4. Vehicle traces parsing. For each road the vehicle passed, only the k bounded preceding roads and the next adjacent road for prediction. Where K is the memory length of a model.
5. Build the PST .Based on the passed road segments.
6. When the vehicle moves to road r_c (r_c =current road segment) at time t,and the roads it just passed are $r_1r_2,\ldots.r_k(k<K)$ .To predict its next road it checks PST was constructed for r_c for corresponding time period .
7. If no such PST was found vehicle is fail to predict the next road.
8. If PST found it searches the next road segments with the longest suffix matching with $r_cr_1r_2 \ldots\ldots\ldots\ldots r_k$ and ending with r_n. where r_n is the one of the unidirectional adjacent road of r_c.

It returns the patterns with highest probability .

It is useful for real time traffic conditions.

5. Conclusion

As we have studied the statistical models for route prediction in previous section .For each model we have to give a trained vehicle route set as input then only we can predict the future road segment. In Simple or hidden Markov model (HMM) to predict the short-term driving route of vehicles, on the static route selection that took no dynamic traffic conditions into account. 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 learn ability hardness. Where using VMM which deploys PST to generate the mobility patterns for long term route prediction, which can be applied for real time traffic conditions for predicting the future route for the driver.

The table1 shows the comparative study of these three models.

The above mentioned all models are having centralized approach, when traffic is so high or more vehicles are involved in it will produce the poor scalability.

References

Heddebaut, M., Rioult, J., Ghys, J.P., Gransart, C., and Ambellouis, S. (2005), "Broadband vehicle-to-vehicle communication using an extended autonomous cruise control sensor," *Meas. Sci. Technol.*, vol. 16, no. 6, pp. 1363–1373.

Shulman, M. and Deering, R. (2005), "Third annual report of the crash avoidance metrics partnership April 2003–March 2004," Nat. Highw. Traffic Safety Admin. (NHTSA), Washington, DC. DOT HS 809 837.

Andreone, L., and Ricerche, C. (2005) , "Activities and applications of the vehicle to vehicle and vehicle to infrastructure communication to enhance road safety," in Proc. 5th Eur. Congr. Exhib. ITS, Hannover, Germany.

Kruger, R., Fuler, H., Torrent-Moreno, M., Transier, M., Hartenstein, H., and Effelsberg, W. (2005) ,"Statistical analysis of the FleetNet highway movement patterns," University of Mannheim, Mannheim, Germany, Tech. Rep. TR-2005-004.

Guangtao Xue, Zhongwei Li, and Hongzi Zhu, (2009), "Traffic-Known Urban Vehicular Route Prediction Based on Partial Mobility Patterns", IEEE.

SAE (2008) World Congress, April-2008, USA."A Markov Model For Driver Turn Prediction".

Ried Simmons, Brett Browning, Yilu Zhang & Varsha Sadekar, "Learning To Predict Driver Route And destination Intent."

Shanugan, K.S and A.M. Breipohl, (1998), "Random Signals: Detection, Estimation and Data Analysis", New York: John Willey & Sons.664.

Dana Ron ,Yoram singer and Naftali Tishby(1996),"The Power of Amnesia: Learning Probabilistic Automata with Variable Memory Length. Machine Learning ".pp117-149.

Fig.1: Vehicular Ad Hoc Network Overview

Table 1: Comparison of statistical models

Markov Model		**Hidden Markov Model**	**Variable Order Markov Model**
1.	It observes only fixed no. of random variables	It observes Hidden states, as well as fixed no. of random variables.	It observes a variable order variables not fixed.
2.	Static approach	Static approach	Dynamic approach
3.	Predict only next route segment dependent on previous road segment.	Predicts next route segment independent on previous road segment.	Generates the sequences of road segments for prediction.
4.	It Doesn't consider other parameters.	It Consider other parameters like Time, Day, Month, Year.	It Consider other parameters like Time, Day, Month, Year.