ZOOM: Scaling the Mobility for Fast Opportunistic Forwarding in Vehicular Networks

Hongzi Zhu[1], Mianxiong Dong[2], Shan Chang[3], Yanmin Zhu[1], Minglu Li[1], Xuemin (Sherman) Shen[4]

[1] Shanghai Jiao Tong University, China
[2] University of Aizu, Japan
[3]Dong Hua University, China
[4] University of Waterloo, Canada

{hongzi, yzhu, li-ml}@cs.sjtu.edu.cn, dongmx@u-aizu.ac.jp, changshan@dhu.edu.cn, xshen@bbcr.uwaterloo.ca

***Abstract*—Vehicular networks consist of highly mobile vehicles communications, where connectivity is intermittent. Due to the distributed and highly dynamic nature of vehicular network, to minimize the end-to-end delay and the network traffic at the same time in data forwarding is very hard. Heuristic algorithms utilizing either contact-level or social-level scale of vehicular mobility have only one-sided view of the network and therefore are not optimal. In this paper, by analyzing three large sets of Global Positioning System (GPS) trace of more than ten thousand public vehicles, we find that pairwise contacts have strong temporal correlation. Furthermore, the contact graph of vehicles presents complex structure when aggregating the underlying contacts. In understanding the impact of both levels of mobility to the data forwarding, we propose an innovative scheme, named ZOOM, for fast opportunistic forwarding in vehicular networks, which automatically choose the most appropriate mobility information when deciding next data-relays in order to minimize the end-to-end delay while reducing the network traffic. Extensive trace-driven simulations demonstrate the efficacy of ZOOM design. On average, ZOOM can improve 30% performance gain comparing to the state-of-art algorithms.**

Keywords-mobility scale; vehicular networks; opportunistic forwarding; social network analysis; inter-contact time

I. INTRODUCTION

Vehicular networks are emerging as a new landscape of mobile ad hoc networks, aiming to provide a wide spectrum of safety and comfort applications to drivers and passengers. In vehicular networks, vehicles equipped with wireless communication devices can transfer data with each other (inter-vehicle communications) as well as with the roadside infrastructure (vehicle-to-roadside communications). In order to successfully transfer data from a vehicle to another, the vehicle has to first wait until it geographically "meets" other vehicles (contact happens) for data-relay. Data communication, therefore, constantly experiences large delays. Fast data forwarding which refers to minimizing the end-to-end delay and network traffic at the same time in vehicular networks is the cornerstone of a wide variety of applications. For example, real-time road traffic information can be obtained by exchanging local traffic observations among vehicles.

Provisioning fast data forwarding in vehicular networks, however, is quite challenging due to three reasons. First, due to the dynamic characteristics of the network, it is very hard to know future contacts between vehicles. With no such information, routing decision made based on past contact statistics can hardly achieve the optimal. To make things worse, even all future contacts are known, finding the fastest path for a given data traffic load is NP-hard [3]. Second, with the distributed nature of the network, a vehicle can only have partial information of the network, which make a globally optimal solution is very hard, if not impossible. Last, data communications are via wireless channels and therefore the network resource is restrained, which makes solutions like epidemic routing [2] [3] infeasible since they introduce prohibitive network traffic.

In the literature, opportunistic message forwarding protocols in intermittently connected mobile ad hoc networks (MANETs) and delay tolerant networks (DTNs) have been proposed. Most existing opportunistic data forwarding algorithms focus on the *contact-level mobility* of nodes. By collecting and analyzing contacts between two vehicles, such algorithms try to find predictive statistics about this pair of vehicles, such as the frequency and the spatial-temporal distributions of contacts and inter-contact times (ICTs) and then use such information to guide data forwarding. Recently, social network analysis has been proposed as a general and powerful tool to forward data in DTNs. By *aggregating* past observed pair wise contacts into a *social* graph, data forwarding algorithms focusing on *social-level mobility* study the network structure, which push packets towards nodes with more important network positions. In general, contact-based algorithms can be very efficient when forwarding packets between regularly encountered nodes but less effective when no prior contact knowledge is available. In contrast, social-based algorithms leverage the knowledge of network structures to route packets. The rationale of algorithms in this category is based on the "small world" phenomenon [14] in social networks. However, short chains of acquaintances do not mean that the delay for delivering messages between a pair of nodes is necessarily short. Furthermore, it is less efficient when forwarding data between regularly-met vehicles. As a result, there is no existing successful opportunistic data forwarding solution, to the best of our knowledge, for the fast data forwarding problem in vehicular networks.

In this paper, we propose an innovative opportunistic data forwarding scheme, named ZOOM, which elegantly integrates both contact-level and social-level mobility for fast routing. The design of ZOOM is based on two key observations found by analyzing three traces of over ten thousand vehicles. First, we find that consecutive ICTs have strong temporal correlations, which can be utilized to predict future contacts. Second, we also find that contact graphs established by aggregating pairwise contacts represent clear social structures. Inspired by these observations, we first train Markov chains to capture the temporal correlations of pairwise contacts, based on which we infer future contact opportunities. We then use centrality to measure the importance of a vehicle in the contact graph. With mobility information in both levels, when two vehicles encounter, the vehicle with shorter expected ICT with the destination will be chosen as the next relay of a packet. If no such information available, the vehicle which has larger network centrality will act as the next data relay. We conduct extensive trace-driven simulation on all three data sets and the results demonstrate the efficacy of ZOOM design. On average, ZOOM can improve 30% performance gain comparing to the state-of-art algorithms.

We highlight our main contributions in this paper as follows:

- By intensive trace analysis, we find that pairwise contacts have strong temporal correlations and train *k*-th order Markov chains to predict future contacts between a pair of vehicles.
- We also find clear social structures existing in the contact graph established by aggregating contacts, which implies that social level mobility may be used to facilitate data forwarding in vehicular networks.
- We organically integrate vehicular mobility characteristics of both contact- and social-level into solving the fast data forwarding problem and achieve great improvement in terms of minimizing end-to-end delay and the network traffic as well.

The remainder of this paper is organized as follows. Section II presents the related work. In Section III, we describe the characteristics of the GPS trace data. Section IV analyzes the impact of different mobility scales to the data forwarding performance. We elaborate the design of ZOOM in Section V. Section VI describes the methodology to evaluate the performance of our message forwarding algorithm and presents the results. Finally, we give concluding remarks and outline the directions for future work in Section VII.

II. RELATED WORK

In intermittently connected networks, data communications are opportunistic. The mobility characteristics of objects are central to data forwarding performance. Based on granularity at which the underlying node mobility is exploited, we divide existing opportunistic forwarding schemes into following three categories:

Random methods: Without requiring any information, random walks [1] can be used for data-relay. For a random walk, a node randomly selects a neighbor as the next hop to carry a message. Using random walks generates moderate network traffic but tends to have very large end-to-end delay. An extreme case is epidemic routing [2] [3], where a message is flooded in the network. Using epidemic routing can achieve the minimum end-to-end delay and maximum delivery ratio but generates unacceptable network overhead at the same time.

Utilizing contact-level mobility: Algorithms residing in this category extensively investigate microscopic mobility properties and their characteristics of nodes to facility data forwarding. For example, in MaxProp [4], likelihood (probability) that a node will encounter the destination of a packet is estimated and used as the forwarding utility. A recursive process has been deployed in [5] to calculate the minimum end-to-end delivery delay, assuming that the tail distribution of ICTs is exponential and ICTs are independent. S. C. Nelson et al. have proposed an encounter-based routing scheme [6] using the rate of encounter of a node as message relay utility. Observing that successive ICTs have strong temporal correlations, Markov chains [7] [8] have been used to predict future contacts. In this category, a utility function is defined and measured for every other node in the network. If the current message carrier meets a node with a higher utility, the message is forwarded to this node. Algorithms in this category would be very effective when delivering packets to those nodes with which a node has prior contact knowledge.

Utilizing social-level mobility: In this category, macroscopic structures of node mobility are characterized by data forwarding algorithms. Pairwise contacts are aggregated to social graphs that reflect the regular social relationships between nodes. For example, E. M. Daly et al. 0 have proposed a social based routing scheme called SimBet, which assesses similarity and betweenness centrality Packets are routed to most central nodes until a node with higher similarity is met. Then the packet is routing within the community until the destination is reached. P. Hui et al. [10] have proposed a similar social based data forwarding scheme called Bubble Rap, where betweenness centrality is also used to find bridging nodes and communities are explicitly identified by a distributed community detection algorithm. J. Pujol et al. [11] have proposed a forwarding algorithm called FairRoute leveraging two social processes called perceived interaction strength and assortativity to distribute load more evenly among nodes in the network. Recognizing the importance of capturing real social relationships to the performance of data forwarding, T. Hossmann et al. [12] have proposed an online algorithm to infer the optimal aggregation density. In this paper, we propose ZOOM which elegantly manages to capture both contact-level and social-level mobility of vehicles in an integrated approach.

III. EMPIRICAL MOBILITY ANALYSIS

A. Collecting Urban Vehicular Trace Data

In order to understand vehicular mobility and conduct informed design of message forwarding algorithms between vehicles, it is of great importance to study the empirical data in terms of frequency, duration and temporal distribution of contacts among them. For this purpose, we use three datasets consisting of traces from two metropolises in China and two types of vehicles, i.e., buses and taxies. Key statistics of the traces are listed in Table I.

Shanghai Buses: The trace consists of GPS reports sent from 2,501 buses which serve on 100 routes and cover the whole downtown area in Shanghai between Feb. 19 and Mar. 5, 2007. A commuting bus periodically sends GPS reports back to a backend data center via GPRS channel. The specific information contained in such a report includes: ID, the longitude and latitude coordinates of the current location, timestamp, moving speed, and heading direction. In addition, other status information, such as whether the bus is arriving at a stop or a terminal is also sent. Due to the GPRS communication cost for data transmission, reports are sent at a granularity of around one minute.

Shanghai Taxies: We also collected the GPS trace of taxies in Shanghai collected between Feb. 1 and Mar. 3, 2007. We chose 2,109 taxies in the datasets which have consecutive GPS reports on each day during the 31 days. The information contained in a taxi GPS report is similar to that of bus except that taxies also report whether passengers are onboard. The granularity of reports is one minute for taxies with passengers and about 15 seconds for vacant ones.

Shenzhen Taxies: The trace collection of taxies in Shenzhen is similar to Shanghai taxi trace. We use the whole month trace in October, 2009. We chose 8,291 taxies which continuously send GPS reports during the whole period. Taxies in Shenzhen always send GPS reports on every one minute.

We choose taxies and buses for the study for two reasons. First, taxies and buses shows two distinct mobility patterns, namely, rather random and well scheduled, respectively. Second, the privacy problem is less concerned since they are public vehicles.

B. Microscope Mobility of Pairwise Contacts

With short-range wireless communication, potential communication opportunities occur only when two vehicles geographically encounter each other. Therefore, contacts are the smallest scale reflecting the mobility of vehicles with regard to data forwarding. In this subsection, we first study the vehicular mobility at contact level.

TABLE I. COMPARISON OF THERE DATA SETS

Data Set	*Shanghai Bus*	*Shanghai Taxi*	*Shenzhen Taxi*
Number of vehicles	2,501	2,109	8,291
From date	Feb. 19, 2007	Feb. 1, 2007	Oct. 1, 2009
Duration (day)	15	31	31
Granularity (second)	60	15*, 60**	60
Number of contacts	1,229,380	22,053,178	23,968,860
Mean ICT (minute)	31.8	47.6	30.5
#Communities	29	56	43
Q	0.8733	0.8471	0.6230

*vacant, **passengers onboard

1) Extracting Pairwise Contacts and Inter-Contact Times

We assume that two vehicles would have a connection opportunity (called a *contact*) if they report their locations within certain period of time and the distance between the reported locations are within the communication range. We use a sliding time window to check contacts between a pair of taxies[1]. Note that, despite the inaccuracy may be introduced by this assumption and the contact extraction algorithm, the essential vehicular mobility characteristics are preserved and therefore the results are very valuable for study.

We refer to an ICT as the time elapsed between two successive contacts of the same vehicles [7] [13]. Specifically, the inter-contact time is computed at the end of each contact, as the time period between the end of this contact and the start of the next contact between the same two vehicles[2]. The key characteristics of contacts and ICTs of all traces are shown in Table I.

2) Mining Temporal Correlations of Successive ICTs

We examine whether the pairwise ICTs emerge under certain pattern in the time dimension. We examine the correlation between ICTs by computing the marginal and conditional entropy known the last ICT.

Let X be the random variable representing the ICTs between a pair of vehicles. If we have observed N ICTs, these ICTs can be presented by a vector $T = (t_0, t_1, \cdots, t_{N-1})$ where t_i, $0 \le i \le N-1$ denotes the i^{th} ICT. The probability of the ICT being j can be computed as x_j/N, where x_j represents the number of ICT being j. Therefore, the entropy of T is:

$$H(X) = \sum_{j=0}^{\infty}(x_j/N)\log_2 \frac{1}{x_j/N}. \quad (1)$$

Let X' be the random variable for the last ICT between this pair of vehicles given the ICT X. X' and X have the same distribution when N is large enough. The vector T can be written as $Q = \{(t_i, t_{i+1}): 0 \le i \le N-2\}$. Therefore, the joint entropy of X' and X can be computed as:

$$H(X', X) = \sum_{(x',x)\in Q} P(x', x)\log_2 \frac{1}{P(x',x)}, \quad (2)$$

where $P(x', x)$ is the number of times (x', x) appearing in Q divided by the total number of elements in Q. With $H(X)$ and $H(X', X)$, the conditional entropy of X given X' is:

$$H(X|X') = H(X', X) - H(X') = H(X', X) - H(X). \quad (3)$$

The cumulative distribution functions (CDF) of the mean entropy and the mean conditional entropy over each pair of vehicles in all datasets are shown in Fig. 1. It can be seen that the conditional entropy is much smaller than the marginal entropy for all types of vehicles. This implies that the uncertainty about the ICT decreases when knowing the last ICT between the same pair of taxies[3]. We conclude that pairwise ICTs have strong temporal correlation.

[1] Short disconnections less than one minute are removed. The detailed contact extraction algorithm can be found in [7].

[2] We do not take into consideration the inter-contact time starting after the last contacts.

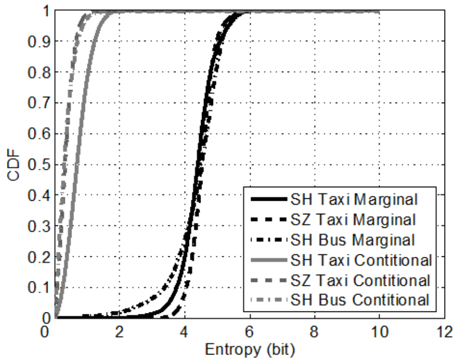

Figure 1. CDF of marginal and conditional entropy of pairwise ICTs between each pair of vehicles in all traces.

C. *Macroscope Mobility of Social Relationship*

Contacts of vehicles actually reflect the complicated social activities of human beings, the characteristic macroscopic structures of human relationships may create complex patterns of contacts, which cannot easily be observed or well understood by only analyzing individual pairwise contacts. For example, people meet "strangers" by chance, "friends" by intention or "familiar strangers" because of their similarity of mobility patterns. In this subsection, we examine the vehicular mobility from a more macroscope perspective.

1) Establishing Contact Graph

We establish a static and weighted contact graph $\mathbb{G}(N, E)$ for each trace by aggregating the entire sequence of contacts between a pair of vehicles. Each vehicle i is a node of the graph, $n_i \epsilon N$, and the edge $e_{ij} \in E$ represents node i and j have certain acquaintance between them. The key to establishing a meaningful contact graph is the metric used to aggregate contacts, which determines whether two nodes share a link and the strength of this connection if exists. Various metrics, such as the number of total contacts observed [10], the age of last contact [15], and the contact frequency and total duration [10], have been used to derive edge strengths. In our study, we use a sliding window to consecutively check the ratio of time with contacts observed to the total period of a trace, called *contact ratio*. There is an edge between two nodes in the contact graph if the contact ratio is higher than a threshold and the weight on this edge takes the contact frequency value. The main reason that we use this metric to aggregate contacts is to reduce the influence of random (*unexpected*) contacts in vehicular networks and comprise as many "regular" relationships as possible. Fig. 2 illustrates the contact graph established on Shanghai taxi trace with sliding window size of one day and contact ratio equal to 60%.

2) Revealing Social Structures

We study the social properties of the contact graph of each trace and examine the degree distributions. The degree of a node in the contact graph is the number of edges incident on this node. We define p_k to be the fraction of nodes in the contact graph that have degree k and plot the complementary cumulative distribution function (CCDF) $P_k = \sum_{k'=k}^{\infty} p_{k'}$. Fig. 3 shows the CCDF of vehicle degree on all traces under semi-logarithmic scale. It is clear to see that all degree distributions have exponential tails,

$$P_k = \sum_{k'=k}^{\infty} p_{k'} \sim \sum_{k'=k}^{\infty} e^{-\frac{k'}{\alpha}} \sim e^{-\frac{k}{\alpha}}. \quad (4)$$

Similar degree distributions have been seen with different networks such as the power grid and railway networks [18]. In contrast, random graphs, where each edge is present or absent with equal probability, have binomial (Poisson in the limit of large graph size) degree distributions.

We further check whether there are *communities* embedded in a contact graph. A community is defined as a subset of nodes with stronger connections between them than towards other nodes, which generally implies a social group. The *modularity* [19] can be used to evaluate the partition of nodes to communities, which is defined as

$$Q = \frac{1}{2m} \sum_{ij} \sum_{r} \left(A_{ij} - \frac{k_i k_j}{2m} \right) S_{ir} S_{jr}, \quad (5)$$

where m is the total number of edges, A_{ij} is the element of the adjacency matrix (if there is an edge between node i and j, $A_{ij} = 1$; otherwise, $A_{ij} = 0$), k_i and k_j are the degree of node i and j ,respectively, and $S_{ir} = 1$ if node i belongs to group r and zero otherwise. Finding the optimal community structure for a contact graph in terms of maximal modularity is an NP-complete problem. We use the Louvain algorithm [20] which iteratively moves each node to an existing community and merges two communities if doing so can maximize the modularity. We choose this algorithm because it has been reported to be fast and has good or better community partition comparing with other algorithms on a different number of graphs [20].

The modularity and the number of found communities for all traces are listed in Table I. From the list, we have the following observations: (1) the modularity values vary in traces but overall are quite high. This implies that urban vehicular networks are highly structured rather than randomly connected ($Q = 0$). High modularity ($Q > 0.3$) can also be seen in other social and biological networks [19]. (2) Buses have higher modularity than taxies. This is easy to understand

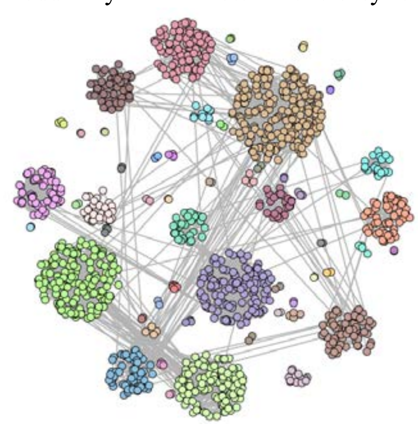

Figure 2. Contact Graph of Shanghai Taxi Trace containing 1,226 nodes, which is highly structured with 56 communities.

[3]Conditional entropy of ICT keeps decreasing as more previous ICTs are known, especially for vehicles constantly meeting with each other. Due to the page limitation, we only show the results when only last ICT is known.

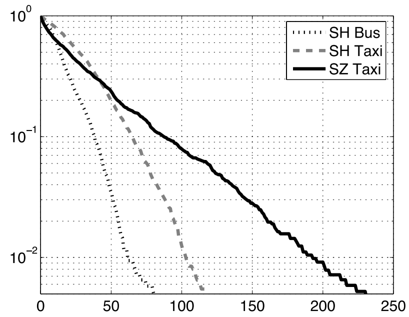

Figure 3. The CCDF of the vehicle degree on all traces under semi-logarithmic scale.

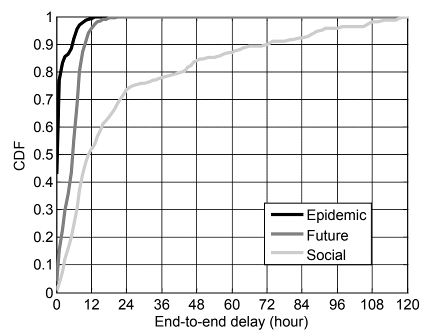

Figure 4. The CDF of end-to-end delay over 1,000 random generated messages using different algorithms.

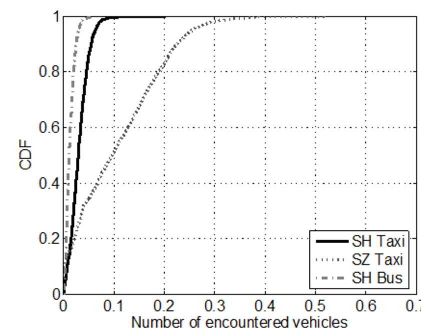

Figure 5. The CDF of the number of encountered vehicles on all datasets.

since buses have dedicated routes and schedules, which makes contacts constant and stable.

IV. Impact of Mobility on Routing Algorithms

In the store-carry-and-forward scenario, the performance of a particular opportunistic forwarding algorithm heavily relies on its capability to accurately capture the underlying mobility of vehicles. In this section, we discuss the impact of different mobility scales to the performance of routing algorithms.

A. *Algorithms Utilizing Contact-Level Mobility*

By collecting and analyzing contacts between two vehicles, it is possible to obtain detailed knowledge about this pair of vehicles such as the contact frequency [4] and the expected delay [7] [8]. Such local knowledge can be used to determine data-relays for a routing algorithm.

To illustrate the performance of algorithms utilizing contact-level mobility of vehicles, we examine a greedy algorithm, called *Future*, in which all future contacts between vehicles are known. In *Future*, a vehicle with messages always chooses a neighboring vehicle which has the shortest delay with the destination. We randomly select 1,000 pair of vehicles as the source and destination of 1,000 messages, using Shanghai taxi trace. Fig. 4 shows the CDF of end-to-end delay over all messages using *Future* and Epidemic routing [3] [4], where a vehicle always forwards its messages to any vehicle it meets. It can be seen that *Future* performs well but experiences larger end-to-end delay comparing to Epidemic routing. For example, above 90% messages can be delivered within six hours using Epidemic routing whereas Future can only reach about 60%.

The main reason for *Future* being sub-optimal is that, without global contact information of other vehicles, *Future* may only find local optimal routing path. Moreover, most vehicles, due to limited mobility, only have contacts with a small portion of other vehicles. For example, Fig. 5 plots the CDF of the ratio of the number of vehicles met by a vehicle to the total number of vehicles in all traces. It can be seen that most Shanghai taxies can only "see" 10% of all taxies. For Shanghai buses, the proportion reduces to about 5% due to the limitation of fixed itineraries and schedules of buses. Comparing to Shanghai taxies, Shenzhen taxies have higher proportion of encountered taxies. The reason seems to be that Shanghai City has larger area than Shenzhen City (three times bigger). Given the same mobility of a taxi, Shenzhen taxies have more opportunities to meet other taxies. Nevertheless, the proportion is still low (e.g., 80% taxies only "see" 20% other taxies).

The consequence of limited view about the whole network is that when a vehicle v_1 is requested to deliver a message to vehicle v_d, it is very likely that v_1 has no knowledge about v_d. To make things worse, when v_1 encounters another vehicle v_2, very likely, v_2 has no information about v_d either. In that case, v_1 has to carry the message until it meets v_d or another vehicle which knows v_d. This will increases the end-to-end delay.

In summary, data forwarding algorithms based on contact-level mobility are effective when delivering packets among "familiar" vehicles with prior contact knowledge but less efficient for "stranger" vehicles.

B. *Algorithms Utilizing Social-Level Mobility*

With contact graph and the social structure observed in the contact graph as described in Subsection III C, a data forwarding algorithm can utilize the social features of nodes or the network to facilitate data forwarding. For example, a greedy hill-climbing procedure is conducted in the network, seeking for more "central" or "popular" nodes in the graph using social network analysis metrics (e.g., centrality and similarity) as data carriers 0 [10]. The rationale of such data forwarding algorithms is based on the "small world" phenomenon in social networks which comes from the observation that individuals are often linked by a short chain of acquaintances (e.g., "six degrees of separation" [14]).

In vehicular network scenario, however, the process of seeking for central nodes as data-relays does not match the goal of the fast opportunistic forwarding problem. A hop in the short paths in social networks may actually undergo a tremendous delay, which is prohibitive for fast data forwarding. In the extreme case, a vehicle can hold a message until it finally meets the destination of the message, which is optimal in terms of the number of hops required to forward the message but definitely not the optimal for minimizing the end-to-end delay. In order to verify our argument, we also conduct

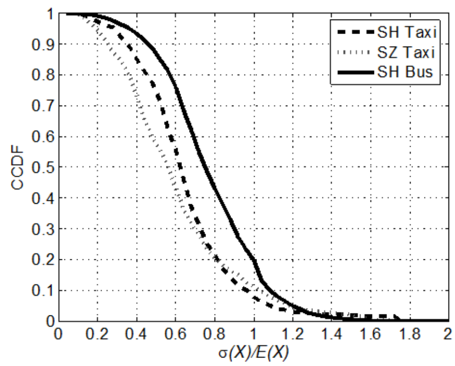

Figure 6. The CCDF of the ratio of standard deviation to the average of contact distribution in a discrete time slot.

an experiment using the same setting as the one described in the above subsection. We evaluate the SimBet algorithm 0 with each node knowing its global social betweenness and similarity values in the contact graph established from Shanghai taxi trace. In SimBet algorithm, a neighboring vehicle is selected as the next relay if a weighted betweenness and similarity utility regard to the destination increases. The CDF of end-to-end delay is shown in Fig. 4. We find that the overall end-to-end delay is quite large. For example, it requires almost three days for 90% messages to be delivered.

The main reason for algorithms utilizing social-level mobility characteristics experiencing large delay is that the social features of vehicles and the network is based on long-term statistics of contacts, which discards the short-term dynamics happening between each pair of vehicles. Specifically, due to the high mobility of vehicles, contacts between two vehicles evolve fast which makes the contact aggregation hard to be accurate. For example, we divide a day into four time slots of six hours and distribute all contacts into respective time slots according to the time when a contact happened. Let X be the random variable that represents the number of contacts exists in a time slot. We then plot the CCDF of the ratio of the standard deviation $\sigma(X)$ to the mean value $E(X)$ for all pairs of vehicles in each trace in Fig. 6. It can be seen that contacts occur quite uneven during a day. For example, for Shenzhen taxies, the probability that the number of contacts between a pair of vehicles can vary 60% comparing to their average number of contacts in a day is above 80%.

Furthermore, without specific contact-level mobility characteristics, social-based algorithms perform less efficient when routing messages among vehicles with acquaintances.

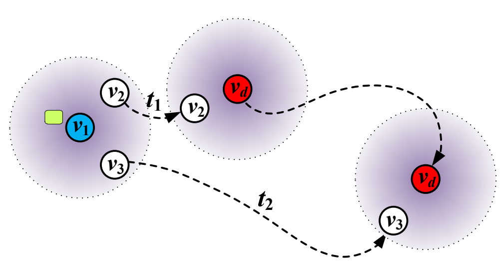

Figure 7. Opportunistic data forwarding scenario, where dashed arrow lines denote the trajectories of vehicles and the disk shades denote the wireless communication range.

For example in Fig. 7, suppose that v_1 has a packet for v_d and encounters v_2 and v_3 at the same time. If v_3 is more "central" than v_2 in the network, v_1 will forward the packet to v_3 even if v_2 will meet v_d sooner than v_3 (i.e., $t_1 < t_2$).

In summary, algorithms based on social-level mobility are effective especially when delivering packets to those "stranger" vehicles but less effective due to the lack of detailed contact-level mobility information.

V. Design of ZOOM

A. Design Overview

From the analysis above, an ideal opportunistic forwarding algorithm should take both contact-level and social-level mobility into account. To this end, we design an innovative opportunistic data forwarding algorithm, ZOOM, which elegantly manages to capture two levels of vehicular mobility in an integrated approach. The core idea of ZOOM is for each vehicle to locally maintain a list of recent contacts with each other encountered vehicle. With the list of past contacts, a vehicle first trains a k-order Markov chain for each other vehicle which can be used to predict the next contact with that vehicle. In addition, it also assesses its position in the network using ego betweenness centrality based on its ego contact graph aggregated from all its contact lists. With the knowledge of the predicted future contact and the ego betweenness, when two vehicles meet, a vehicle carrying a packet first compares its predicted contact delay with the destination of the packet with that of the other vehicle. The vehicle with shorter contact delay estimation will act as the next data-relay. If both vehicles have no contact predictions with the destination, the vehicle having more important position in the network is chosen to carry the packet.

In the following subsections, we first describe our method to capture the contact-level mobility using k-order Markov models to predict future contacts. Then we present the techniques to establish social-level mobility by aggregating fast evolving contacts and calculating the network position of vehicles using ego betweenness. Finally, we describe the opportunistic forwarding strategy of ZOOM.

B. Predicting Contact-Level Mobility

With the strong temporal correlations of successive ICTs embedded in vehicular mobility as described in Section III, we predict ICTs using Markov chains of k-th order [7].

More specifically, let $\{x_i\}_{i=1}^{n}$ be an observed sequence of ICTs between this pair of vehicles. The k-order state transition probabilities of the Markov chain can be estimated for all $a \in \mathcal{S}$ and $\underline{b} \in \mathcal{S}^k$, $\underline{b} = (b_1, b_2, \cdots, b_k)$ as follows. Let $n_{\underline{b}a}$ be the number of times that state $\underline{b}$ is followed by value a in the sample sequence. Let $n_{\underline{b}}$ be the number of times that state $\underline{b}$ is seen and let $p_{\underline{b};a}$ denote the estimation of the state transition probability from state $\underline{b}$ to state $(b_2, \cdots, b_k, a)$. The maximum likelihood estimators of the state transition probabilities of the k-th order Markov chain are

$$p_{\underline{b};a} = \begin{cases} n_{\underline{b}a}/n_{\underline{b}}, & if\ n_{\underline{b}} > 0 \\ 0, & otherwise \end{cases}. \quad (6)$$

C. Establishing Social-level Mobility

To utilize social-level mobility to facilitate opportunistic data forwarding in vehicle networks, ZOOM has to deal with two challenges. One is to accurately aggregate high-dynamic contacts so that the real social-level mobility can be reflected in the contact graph. The other one is to accurately assess the importance of an individual vehicle in the network without the global information.

1) Aggregating Evolving Contacts

We use the aggregation method introduced in Subsection III C. We use six hours as the best sliding time window for contact aggregation as it is reported in [7] that the redundancy of ICTs reaches the minimum when the time difference between two small sets of consecutive ICTs equals to six hours. This also implies that the maximum mobility diversity can be observed within six hours. Increasing the size of the sliding window will reduce the mobility diversity which degrades the accuracy of contact aggregation. It is important to note that different and more sophisticated aggregation schemes are possible, such as online algorithms [12]. Our goal here is to demonstrate that capturing social-level mobility as a complementary counterpart of contact-level priors can significantly improve the performance of opportunistic data forwarding.

2) Calculating Centrality with Local Information

Centrality in graph theory and network analysis is a quantification of the relative importance of a vertex in the graph. It is a nature measure of the structural importance of a node in the network.

In ZOOM, we use *betweenness* [16] to measure the centrality of vehicles, which refers to the extent to which a vehicle lies on the social paths linking other vehicles. Therefore, a vehicle with a high betweenness has a capability to facilitate interactions between the vehicles it links. With only local information, we adopt the algorithm [17] to calculate the betweenness in *ego networks*, which refers to a network consisting of a single vehicle (ego) together with the vehicles (alters) the ego is connected to and all the links among those vehicles. Although the betweenness in ego networks does not correspond perfectly to the global betweenness, the ranking of vehicles are identical in the network. Mathematically, we present the relationships between an ego vehicle v_i and its neighbors in the ego network by a $m \times m$ symmetric matrix A,

$$A_{ij} = \begin{cases} \theta_{i,j}, & if\ \theta_{i,j} > 0 \\ 0, & otherwise \end{cases}. \quad (7)$$

where m is the number of neighbors and $\theta_{i,j}$ is the regularity ratio between v_i and v_j. The ego betweenness of v_i can be calculated as the sum of the reciprocals of the entries of $A^2[1-A]_{i,j}$ [17]. The ego betweenness of a vehicle is updated upon each contact with other vehicles. Specifically, when vehicle v_1 meets v_2, v_2 sends a list of neighbors in its ego network to v_1. Upon receiving the neighbor list, v_1 checks each neighbor in the list, $v_i, i \in [1,2,\cdots,l]$, if v_i is also a neighbor of v_1, then elements $A_{2,i}$ and $A_{i,2}$ are set to $\theta_{2,i}$. If v_2 is a newly encountered vehicle, v_1 will first enlarge $A_{m\times m}$ to $A_{(m+1)\times(m+1)}$ by inserting a new row and a new column for v_2. Then it performs the ego betweenness calculation accordingly. Vehicle v_2 conducts the same operations as v_1 at the same time.

D. Opportunistic Forwarding Strategy

In ZOOM, when vehicle v_1 encounters v_2, v_2 will send a list of all its neighbors and a list of destinations of packets it is currently carrying to v_1. Vehicle v_1 then update the Markov chain and calculates its ego betweenness. For the destination v_d of a packet of v_2, let $\underline{b}_{v_1,v_d}$ denote the current state in the k-th order Markov chain between v_1 and v_d. The estimated delay of the next contact between v_1 and v_d, $\mathcal{E}_{delay}^{v_1,v_d}$ can be calculated as,

$$\mathcal{E}_{delay}^{v_1,v_d} = \sum_{a=0}^{\lfloor \mathbb{T}/\lambda \rfloor} p_{\underline{b}_{v_1,v_d};a} \cdot a. \quad (8)$$

Vehicle v_1 will act as the next relay for this packet if one of the three following cases happens: 1) v_1 is the destination of this packet, i.e., $v_1 = v_d$; 2) v_1 has a shorter estimated delay of the next contact between v_1 and v_d than that between v_2 and v_d, i.e., $\mathcal{E}_{delay}^{v_1,v_d} < \mathcal{E}_{delay}^{v_2,v_d}$ and 3) both v_1 and v_2 have no prior about v_d and v_1 has a larger betweenness value than v_2. After transmitting the packet to v_1, v_2 removes this message from its buffer. Similarly, v_2 conducts the same operations accordingly.

VI. Performance Evaluation

A. Methodology

In this section, we compare our opportunistic forwarding algorithm with several alternative schemes:

- **Epidemic.** In this scheme [2] [3], vehicles exchange every packet whenever they experience a contact. If vehicles have infinite buffer size, using epidemic routing will find the shortest path between the source and destination vehicles and therefore has the shortest end-to-end delay. On the other hand, since there is no control on data forwarding, it also generates a tremendously large volume of network traffic, overwhelming limited wireless bandwidth.
- **Markov.** This scheme [7] establishes a kth order Markov chain to predict the time when the next contact may occur between a pair of vehicles, utilizing the temporal correlations of consecutive ICTs. A greedy strategy is taken in making routing decisions where the neighboring vehicle with the least estimated meeting time with the destination will be chosen as the next data relay.
- **SimBet.** This scheme 0 assesses similarity between nodes in a social graph to detect nodes residing in the same community, and betweenness centrality to identify

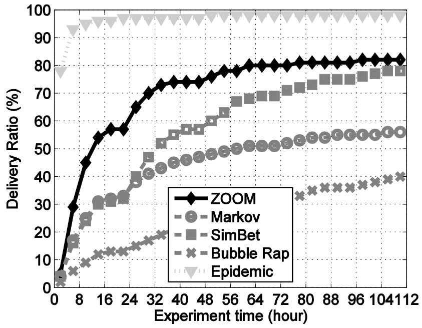

Figure 8. The average delivery ratio vs. the experiment time.

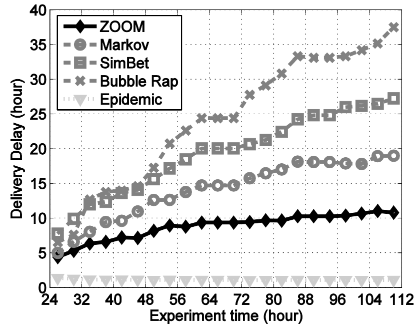

Figure 9. The average delivery delay vs. the experiment time.

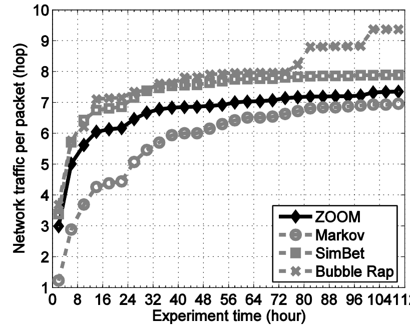

Figure 10. The network traffic vs. the experiment time.

bridging nodes which could carry a packet from one community to another. Packets are routed to the most central nodes until a node with higher similarity with the destination is met. Then the packet is forwarded within the community until the destination is reached.

- **Bubble Rap.** This scheme [10] uses a similar approach as SimBet except that communities here are explicitly identified by a detection algorithm.

We consider four important metrics to evaluate the performance of ZOOM and the above schemes:

1) Delivery ratio. It refers to the ratio of successfully delivered packets to the total number of packets at the end of an experiment.

2) End-to-end delay. It refers to the delay for a packet to be received at its destination. We only calculate end-to-end delay for successfully delivered packets.

3) Network traffic per packet. It refers to the average network cost per packet, calculated by dividing the total number of data forwarding hops by the total number of packets.

4) Packet utility. It refers to the average benefit in reducing the delivery delay by each forwarding hop, calculated by dividing the total amount of time saved (i.e., the time period starting since a packet is delivered and ending when the experiment ends) for all packets to the total number of data forwarding hops.

In the following simulations, we evaluate the above metrics of ZOOM, using real trace data of Shanghai buses for demonstration. We randomly choose 1,000 buses, and use the contact records of three weeks from Feb. 19 to Feb. 28, 2007 for the initialization of all alternative schemes and use contact records of four and a half days from 8am on Mar. 1 to Mar. 5, 2007 for data forwarding experiments (the reason that we set the experiment to start from 8am in the morning is because most buses are not in service at night.). At the beginning of each experiment, we inject 100 packets using a Poisson packet generator with a mean interval of ten seconds. For each packet, the source and destination are randomly chosen among all buses in the data set. Here we make a general assumption that two vehicles can always successfully conduct all data transmission when they have a contact. We run each experiment 50 times and get the average.

B. Performance Comparison

In this simulation scenario, we compare ZOOM with all other alternative forwarding algorithms. For the sake of fairness, we adjust the contact aggregation scale for the best delivery performance for SimBet and Bubble Rap. In this simulation setting, the optimal number of contacts for a pair of vehicles to have a link in the contact graph is twenty.

Fig. 8 plots the average delivery ratio as a function of experiment time. It can be seen that ZOOM outperforms other algorithms except the epidemic routing. As epidemic routing can always find the shortest path by aggressively spreading a packet over the whole network, it also causes unacceptable network traffic. It can be seen that ZOOM is capable of obtaining great delivery ratio gain in a very short period of time. For instance, ZOOM can successfully deliver over 60% packets within 24 hours while the ratio for Markov, SimBet and Bubble Rap is 35%, 37% and 24%, respectively. In addition, it is very interesting to see that, for all schemes, the delivery ratio stabilizes and stops to increase when it is night, for example, during the first night from the 14^{th} hour (i.e., 10pm on Mar. 1) to the 22^{th} hour (i.e., 6am on Mar. 2), and the second night from the 38^{th} hour (i.e., 10pm on Mar. 2) to the 46^{th} hour (i.e., 6am on Mar. 3). The reason is that the data forwarding process would suspend during the night as most buses are not in service at night and would continue in daytime when buses are on duty.

Fig. 9 and Fig. 10 plots the average end-to-end delay and network traffic as a function of experiment time for all successfully delivered packets, respectively. Note that, for comparison fairness, we only take into account those packets that can be successfully delivered by all schemes. It is clear to see that, in general, algorithms utilizing contact-level mobility can achieve very small delivery delay comparing with social-level-mobility-based routing algorithms. Moreover, ZOOM can achieve the minimum end-to-end delay (excluding the epidemic routing). It can be seen that Markov generates the least network traffic and ZOOM introduces slightly more traffic than Markov. Combining both the end-to-end delivery delay and the network traffic, we argue that ZOOM can actively spend few more hops to achieve far more gain in end-to-end delivery delay. In contrast, schemes based on social-level mobility spend more hops but result in larger delays and, therefore, have less network-cost-efficiency.

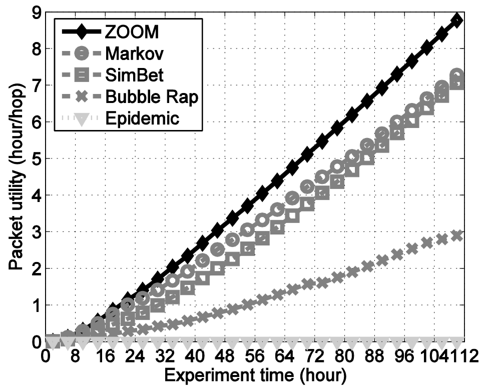

Figure 11. The packet utility vs. the experiment time.

In general, routing algorithms try to trade off between delivery delay and network traffic cost. Short delivery delays usually imply large network traffic. To better measure how efficient a routing algorithm can be, we evaluate all schemes with the packet utility metric. Fig. 11 plots the average packet utility as a function of experiment time. It is clear that ZOOM has the highest packet utility among all schemes. In summary, ZOOM is a very fast and cost-efficient opportunistic routing scheme under urban VANET settings.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an opportunistic forwarding algorithm ZOOM which captures both lower level mobility at pairwise contact scale and upper level mobility from the VANET perspective. Our algorithm uses locally collected contacts to predict the future contact opportunities between vehicles. Moreover, the capability to predict contacts is then utilized to reflect the social relation ties between vehicles. We have demonstrated the efficacy of our algorithm through extensive trace-driven simulations. For our future work, we intend to look into various traces, as well as the realistic vehicular mobility, in order to better understand the underlying structure and similarity of vehicles. Furthermore, we will explore more sophisticated mappings such as appropriate weighted graphs. In addition, we will also investigate more accurate calculation method for the packet utility metric.

ACKNOWLEDGEMENT

This research was supported in part by National Natural Science Foundation of China (Grants No. 61202375, 60970106, 61170237, 61170238, 60903190, 61027009), the National High Technology Research and Development Program (2011AA010502, 2011AA010101), Doctoral Fund of Ministry of Education of China (20100073120021), STCSM Grant 12ZR1414900, Shanghai Pu Jiang Talents Program (10PJ1405800), Shanghai Chen Guang Program (10CG11), JSPS A3 Foresight Program.

REFERENCES

[1] C. Gkantsidis, M. Mihail, and A. Saberi, "Random walks in peer-to-peer networks," *in Proceeding of IEEE INFOCOM* 2004, Hong Kong, China, Mar, 2004.

[2] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Efficient routing in intermittently connected networks: the multi-copy case," *ACM/IEEE Transactions on Networking*, vol. 16, no. 1, pp. 77-90, 2008.

[3] A. Balasubramanian, B.N. Levine, and A. Venkataramani, "DTN routing as a resource allocation rroblem," *in Proceeding of ACM SIGCOMM 2007*, Kyoto, Japan, pp. 372-384, Aug. 2007.

[4] J. Burgess, B. Gallager, D. Jensen, and B. Levine, "MaxProp: routing for vehicle-based disruption-tolerant networks," *in Proceeding of IEEE INFOCOM* 2006, Barcelona, Catalunya, Spain, Apr. 2006.

[5] V. Conan, J. Leguay, and T. Friedman, "Fixed point opportunistic routing in delay tolerant networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 773-782, 2008.

[6] S. C. Nelson, M. Bakht, and R. Kravets, "Encounter-based routing in dtns," *in Proceeding of IEEE INFOCOMM 2009,* Rio de Janeiro, Brazil, pp. 846-854, Apr. 2009.

[7] H. Zhu, S. Chang, M. Li, S. Naik, and X. Shen, "Exploiting temporal dependency for opportunistic forwarding in urban vehicular networks," *in Proceeding of IEEE INFOCOMM 2011*, Shanghai, China, Apr. 2011.

[8] Q. Yuan, I. Cardei, and J. Wu, "Predict and relay: an efficient routing in disruption-tolerant networks," in *Proceeding of ACM MOBIHOC 2009*, New Orleans, USA, May. 2009.

[9] E. M. Daly, and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant manets," *in Proceeding of ACM MOBIHOC 2007*, Montreal, Canada, pp. 32-40, Sep. 2007.

[10] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble Rap: social-based forwarding in delay tolerant networks," *in Proceeding of ACM MOBIHOC 2008*, Hong Kong, China, May. 2008.

[11] J. Pujol, A. Toledo, and P. Rodriguez, "Fair routing in delay tolerant networks," *in Proceeding of IEEE INFOCOM 2009*, Rio de Janeiro, Brazil, pp. 837-845, Apr. 2009.

[12] T. Hossmann, T. Spyropoulos, and F. Legendre, "Know thy neighbor: towards optimal mapping of contacts to social graphs for dtn routing," *in Proceeding of IEEE INFOCOM 2010*, San Diego, USA, Mar. 2010.

[13] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on the design of opportunistic forwarding agorithms," *in Proceeding of IEEE INFOCOM 2006*, Barcelona, Catalunya, Spain, Apr. 2006.

[14] S. Milgram, "The small world problem," *Psychology Today*, vol. 1, no. 1, pp. 61-67, 1976.

[15] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, "Age matters: efficient route discovery in mobile ad hoc networks using encounter ages", in Proc. *ACM MobiHoc*, 2003.

[16] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35-41, 1977.

[17] M. Everett, and S. P. Borgatti, "Ego network betweenness," *Social Networks*, vol. 27, issue 1, pp. 31-38, 2005.

[18] Amaral, L. A. N., Scala, A., Barth ´ el ´ emy, M., and Stanley, H. E., Classes of small-world networks, Proc. Natl. Acad. Sci. USA 97, 11149–11152 (2000).

[19] M. E. J. Newman, "Modularity and community structure in networks", *PNAS*, 2006.

[20] V. D. Blondel, J. L. Guillaune, R. Lanbiotte, and E. Lefebvre, "Fast unfolding the communities in large networks", *J. STAT. MECH.*, 2008. .