Multi-Modal Message Dissemination in Vehicular Ad Hoc Networks

Lei Zhang, Yanyan Zhuang, Jianping Pan, Lovereen Kaur, and Hongzi Zhu

1University of Victoria, BC, Canada; 2Anna University, Chennai, India; 3Shanghai Jiaotong University, China

Abstract—As a fundamental network property, inter-contact time (ICT) determines the critical performance metrics of a vehicular ad hoc network (VANET). In this paper, we extract the contact information from the Global Positioning System (GPS) traces of public and semi-public vehicles, including both buses and taxis, in a large modern city. By exploring the spatial and temporal properties of the contact behaviors between both types of vehicles, we propose a multi-modal VANET message dissemination scheme, by letting buses assist the message dissemination among taxis. This scheme leverages the wide time and space coverage of taxis, and the high route and schedule regularity of buses. With the help of buses, the trace-driven simulation shows that both the end-to-end delay and delivery ratio are significantly improved without much extra overhead introduced, which indicates the great potential of exploring heterogeneous mobility patterns in mobile ad hoc networks.

Index Terms—Inter-contact times, vehicular ad hoc networks, message dissemination delay, network traffic

I. INTRODUCTION

The recent advances in mobile communications have expanded the application regime of wireless networks, from supporting laptops or sensor nodes that are mostly stationary, to a much more challenging environment: mobile ad hoc networks (MANET). With a rich history of data trace measurement studies, mostly for mobile users in wireless LANs [1] and MANET [2], [3], it has been deeply understood that the underlying mobility pattern is the most important factor in determining network performance. Vehicular ad hoc networks (VANET), a major component of the future intelligent transportation systems (ITS), are emerging as a new category of MANET aiming at providing both safety and infotainment applications to drivers and passengers. However, with the high mobility of vehicles, no instantaneous end-to-end connectivity always exists for an arbitrary pair of mobile nodes. As a result, vehicles have to rely on the opportunistic contact with each other for message dissemination.

The message dissemination in VANET happens in a store-carry-and-forward manner, where vehicles exchange data packets when they are within the communication range of each other [4], [5]. The opportunistic vehicular contact behaviors, such as the contact frequency and duration, ultimately determine the fundamental network performance metrics, such as the end-to-end delay and the induced network traffic. The time interval between two consecutive contacts of an arbitrary pair of vehicles, referred to as inter-contact time (ICT), is an important characteristic of vehicular contacts that has been commonly measured and analyzed in the existing work.

With the recent work on the trace measurement and analysis in vehicular networks [6]–[8], more insights have been revealed for the intermittent connectivity between vehicles. By analyzing the ICTs between taxis, in particular, it has been shown that the frequent contacts between vehicles are a desirable feature for message delivery and the design of routing algorithms [9]–[11]. Therefore, it is of great importance to understand the properties of ICTs, for the purpose of accurately characterizing the data transfer opportunities in a highly mobile network environment, such as VANET.

Most existing work focused on a homogeneous network, i.e., all mobile nodes have very similar mobility patterns. For the first time in the literature, both measurement analysis and the design of a new message dissemination scheme are presented in this paper, on a heterogeneous vehicular network that consists of both public (buses) and semi-public vehicles (taxis). Buses run on certain routes with predefined schedules and move in a cyclic fashion, but the service regions are limited to the major roads in a city. Taxis, on the other hand, have much wider time and space coverage, but are much less predictable in mobility than buses. By leveraging such different vehicle mobility patterns in a multi-modal VANET message dissemination scheme, the disadvantage of irregular taxis routes, and the limited time and space coverage of buses will be eliminated. Extensive trace-driven simulation on the new message dissemination scheme shows that, the end-to-end delay is significantly reduced when compared with the schemes in the existing work using taxis only, without incurring too much extra overhead. To achieve the same performance, existing schemes actually have higher overhead. The results also show the great potential of exploring heterogeneous mobility patterns in MANET.

The rest of this paper is organized as follows. Section II gives an overview of the most related work done on VANET trace measurement and analysis. Our measurement analysis are presented in Section III, showing both the ICT regularity at the intra/inter-route level and the ICT behavior between buses and taxis. The results of our trace-driven simulation for the multi-modal message dissemination scheme are given in Section IV. Section V concludes the paper with the future work listed.

II. BACKGROUND AND RELATED WORK

[6] is the first work in the literature which conducted a thorough study on bus contacts. The authors discovered that at the bus route level, the contact behavior is highly regular. However, due to the limited size of the bus network measured, i.e.,
around 40 buses in total, [6] only studied the contact behavior within the same bus route. Recent work [7] utilized bus stops as base stations, which store the messages from traveling buses and later disseminate the message opportunistically to other buses in the network. Although the delivery ratio and delay are improved, the requirement of installing a base station at every bus stop, however, would incur a very high deployment cost. Exploring the bus contacts within one route, and utilizing all bus stops for message forwarding, are the two extreme cases.

In the authors explored the regularity and motion cycles of people and public transportation systems, and proposed a routing algorithm by modeling through a probabilistic time-space graph and applying the Markov decision process.

Apart from the discovery of the power-law contact behavior [12] between human beings in the existing work, [11] found that the inter-contact times between taxis in large cities demonstrate an exponential tail distribution. The conclusion is that vehicles in urban environments tend to meet very frequently. The follow-on work [10] and [9] then utilized this contact property for routing schemes in delay-tolerant networks. However, the models and approaches, such as the Markov scheme proposed in [9], require an excessively long learning period, e.g., three weeks of the measured data. The average end-to-end delay between 20 and 30 hours, is not desirable for most applications, even delay-tolerant ones.

There are a few vehicle traces publicly available, e.g., [13] which is the bus traces used in [6], taxi cab traces in the San Francisco area [14], and the bus traces in Seattle [15]. The traces used in this paper, which are also used in [9]–[11], contain GPS records from both buses and taxis. In contrast to [6]–[8] which only analyzed the bus network, and [9]–[11] whose focus was on taxis, such a feature in the trace gives us a unique opportunity to explore the contact behaviors between heterogeneous vehicles, i.e., both buses and taxis, which is presented for the first time in the literature. In Section III we shall see that the number of vehicles in the trace is large, which also gives us the chance to thoroughly understand such vehicular ad hoc networks using a large set of real-world data.

In [17] and [18], the authors proposed social-based packet forwarding protocols, by using RSUs that are placed in hot spots in cities as message forwarding stations. As a result, no information about the receiver location is needed as long as receivers could pick up their messages from these hot spots. Compared with [17] and [18], our scheme does not need external base stations. Instead, the public transportation system is utilized as a huge virtual base station with good stability and connectivity, potentially preserving receiver privacy as well.

### III. Measurement Analysis

#### A. Inter-Contact Times Extraction

We collect inter-contact time (ICT) statistics from the GPS traces of hundreds of taxis and buses in Shanghai, China. There are in total more than 4,000 taxis and 2,500 buses in the trace data, from which we make a careful selection and sampling (shown below). As in the literature, inter-contact time in this paper is defined as the time elapsed between two consecutive contacts of the same pair of vehicles [6], [9], where two vehicles have a contact opportunity whenever their locations at a certain time are within a given communication range.

As buses and taxis travel in the city, they periodically send GPS reports back to a data center via an on-board GSM device. For taxis, such reports contain the information of the taxi ID, the longitude and latitude coordinates of the current location, timestamp, speed, heading and operational status, with 1 indicating hired and 0 otherwise. For buses, the reports also contain the route ID that the bus is operating on, and whether the bus is at the terminal station, etc. However, due to the GPRS communication cost and data collection errors, the collected reports from each vehicle have a very low granularity: about one minute interval between the consecutive reports from the same taxi, and one to 30 minutes between the consecutive reports from the same bus. It is also found that the traces are often erroneous in certain fields, such as the instantaneous vehicle speed, whereas the longitude and latitude are relatively accurate, if not shadowed by high-rise buildings in the downtown core area.

In order to handle such noisy data, linear interpolation was used to increase the trace granularity. The interpolation method is linear, because vehicles typically travel along straight road segments in a short time duration, and such method also has a lower computational cost. Table I shows the number of vehicles in the traces. Among more than 4,000 taxis, we selected 500 taxis uniformly at random. The uniform selection is based on the fact that the behaviors of different taxis in an urban area are very similar in a statistical manner [16].

The selected bus routes are shown in Fig. 1, where different colors correspond to the bus traces on different routes. The seven routes are located in the urban area of Shanghai, where the major financial and tourism districts, universities and train stations are located. These routes cover the major roads that

<table>
<thead>
<tr>
<th>TABLE I: GPS Trace Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of taxis</td>
</tr>
<tr>
<td>Number of selected taxis</td>
</tr>
<tr>
<td>Total number of buses/routes</td>
</tr>
<tr>
<td>Number of selected buses/routes</td>
</tr>
</tbody>
</table>

Fig. 1: Selected Bus Backbone.
carry most of the traffic in the city, and form a grid-like backbone. On the other hand, taxis have a much wider time and space coverage in both the urban and suburban areas, and operate even when buses are not in service at midnight.

Applying the linear interpolation to the discrete, noisy GPS reports, the updated data granularity is one second, a much higher resolution than that in [9], etc. In each second, the distance between each pair of vehicles is checked, given a communication range of 100 m. Two vehicles have a contact whenever their locations are within this range. The inter-contact time is then measured as the duration of the time from the end of a contact to the beginning of the subsequent contact. Furthermore, for each pair of vehicles, any two subsequent contacts that occur within one minute of each other are combined. This is to reduce the very short inter-contact times when two vehicles that travel closely in the same direction come in and out of the communication range of each other, as their spacing changes due to the road traffic. In this situation, the multiple short contacts are merged into one.

B. Intra-Route ICT

Ideally, the buses on a route are dispatched evenly at a certain time and run at a constant speed, as the triangular traces shown in Fig. 2(a). The y-axis is the distance offset of the bus to one of the terminal stations. Define RTT as the round trip time from the moment when a bus is dispatched from a terminal, till the time it returns to the terminal. The intersection points of different bus trajectories are the contacts occurred between these buses. It is obvious from Fig. 2(a) that the buses on the same route contact each other regularly every half RTT, regardless of the spacing between these buses. In reality, the bus dispatch times are not uniform, and the time when buses stay at a stop or terminal is not fixed. Figure 2(b) plots the real traces of three buses running on the same route, from 8AM to 6PM on a particular day. The markers are the contacts between each of the buses to all others on the same route. From the figure, although the dispatch times and contacts are not as regular as those in the ideal traces, the inter-contact times between the same pair of buses are still around half RTT.

Figure 3(a) and (b) show the cumulative distribution functions (CDF) of ICTs between bus pairs on the same routes. Route 01 and 123 in Fig. 3(a) have very different bus round trip time, i.e., $RTT/2 = 70\text{ min}$ and $37\text{ min}$, respectively. In Fig. 3(b), for all three routes 910, 931 and 985, $RTT/2 \approx 1\text{ hr}$. In both figures, the CDF has a jump at a periodicity of every $RTT/2$, corresponding to the high probability of buses contacting each other at integer multiples of $RTT/2$. The amplitude between these periodical jumps is equal to the peaks of contact probability at these local maximum, the value of which depends on number of buses on the route, the spacing between them, and the length of the route. We can observe a decay in the consecutive amplitudes, indicating the probability for a bus pair to encounter vanishes over the time. These results are similar to those in [6]. From the figures, it is obvious that buses on the same route encounter each other very frequently: more than 60% of buses have contacts with other buses on the same route in less than 2 hours.

C. Inter-Route ICT

Contacts occur between buses on different routes if the two routes have overlapping, as shown in Fig. 4(a). $y$ is the length of the segment that is common to both routes, whereas $x_1$,
\[ ICT = \begin{cases} \frac{T_1}{2} - \frac{i x_1 - j x_2}{v} & \text{if } i = j = 0 \\ \frac{T_2}{2} + \frac{i x_1 - j x_2}{v} & \text{if } i = j = 2 \end{cases} \]

where \( i \) and \( j \) are integer numbers. The combinations of \( i \) and \( j \) have to satisfy that the intersections of bus trajectories, as shown in Fig. 4(a), occur at the common segment of length \( y \).

Figure 4(b) shows the CDF of inter-route ICTs extracted from bus traces. The jumps in CDFs happen at the point where contacts happen at the overlapping region of the two routes. For example, for route 01 and 985 in Fig. 4(b), when \( i = j = 0 \) in (1), the first peak happens at 30 min (indicated by the solid vertical line). The next peak occurs when \( i = j = 2 \) at 122 min (the dash vertical line). The calculation for route 13 and 36 can be done with the same approach. Note that [6] was unable to analyze the inter-route ICT because of a limited number of buses measured.

From Fig. 4(b), selected buses also contact each other frequently even though they run on different routes: more than 60% of inter-route contacts happen within 4 hours. In Section IV, the results of the end-to-end delay from the trace-driven simulation also show that the contacts among the seven selected bus routes are frequent enough to ensure a good stability and connectivity of the whole bus backbone.

D. Bus-Taxi ICT

During a typical day, a taxi travels between different districts of the city. It is highly probable that it encounters the buses on the bus backbone. Similarly, buses run on major roads in the urban districts, which taxis also travel on frequently. Figure 5 shows the aggregated ICTs between any pair of backbone bus and taxi. In metropolis as Shanghai, it is not surprising to see that about 95% of taxis meet a bus on the selected backbone within an hour, and more than 95% of these buses encounter a taxi in less than 20 min. The reason for this asymmetry is that, taxis travel in both urban and suburban areas, where the latter has almost no selected bus running.

The contacts in Fig. 5 happen between heterogeneous vehicles of buses and taxis. From the frequent bus-taxi contact shown in Fig. 5, we can conjecture that if the connectivity within the bus backbone shown in Fig. 1 is good, then the seven selected bus routes can be treated as a virtual base station which can store-carry-and-forward messages between different taxis. The contacts between taxis have been investigated in the existing work [9], including their previous work [10], [11]. However, the work of utilizing heterogeneous vehicular contacts in message dissemination schemes, is presented in the literature for the first time in this paper.

IV. MULTI-MODAL MESSAGE DISSEMINATION WITH TRACE-DRIVEN SIMULATION

In this section, we propose a multi-modal message dissemination scheme, where an arbitrary taxi tries to deliver a message to another taxi. Taxis as message forwards, have the options of forwarding messages to the other taxis, or forwarding to any of the buses on the bus backbone if they encounter each other, i.e., multi-modal message dissemination. The latter is motivated by the fact that taxis are highly likely to meet any of the selected buses within a short time, as shown in Fig. 5. We treat the bus backbone as a virtual base station that opportunistically carries and forwards messages between taxis. Because the good stability and connectivity among the selected buses observed in Fig. 3 and Fig. 4, we can expect the network
delay will be reduced when using buses as the opportunistic relay, if not much extra network overhead introduced.

To fully understand the impact of the bus system on the existing message dissemination schemes among taxis, an oracle model is used. Similar to the learning procedure in [9], the oracle is based on the vehicle contact information history, i.e., a vehicle always forwards the message to another vehicle with a higher probability encountering the destination according to the contact history. Among different message propagation strategies, the oracle-based dissemination always gives a performance upper bound for the existing work, such as [9]. With the public bus system, the network performance metrics in terms of end-to-end delay and delivery ratio can be further improved with a reasonable overhead, compared with the multi-copy message propagation among taxis.

A. Multi-Modal Message Dissemination

In the following, we conduct a trace-driven simulation study and compare the performance of two main schemes: T2T, where taxis only forward messages to other taxis; T2B2T, where taxis could forward messages to buses, and then let buses help relay messages towards the destination taxi. Each of the two schemes has different strategies as explained below: i) Single-copy T2T, where only one copy of the message is disseminated by the source taxi and then propagated through the network; ii) multi-copy T2T, where the source taxi disseminates a copy of the message whenever it has a contact with another taxi, and the other taxi only acts as a forwarder towards the destination. In both i) and ii), the taxis other than the source do not create new copies of the message, and the oracle contact model is used in the message dissemination strategy. In multi-copy T2T, the total number of copies of a message is determined by the number of taxi contacts within the message time-to-live (TTL). It is obvious that multi-copy T2T is more likely to deliver messages with a lower delay, since it creates more forwarding paths for message delivery.

In T2B2T, we assume that buses have sufficient buffer space. Messages are propagated among the buses, and buses relay the messages to taxis according to two strategies: i) un-controlled T2B2T, where the bus keeps the message after being forwarded to a taxi, in order to forward the message to other buses and taxis; ii) controlled T2B2T, i.e., whenever a bus hands over a message to a taxi, the bus deletes the message from its buffer. In both i) and ii), only one copy of message is disseminated from the source taxi. In ii), a message can only enter the bus backbone once, to avoid the unnecessary ping-pong effect that consumes an excessive amount of network resources. The controlled strategy creates a smaller number of message copies, while with the un-controlled strategy, messages are more likely to be delivered to the destination taxi with a lower delay.

The above schemes are verified through the trace-driven simulation, using the contact traces among the 500 taxis and 192 buses outlined in Table I. Taxis start message dissemination at 7AM, the typical starting time of a workday. All the following results are averaged over 200 simulation runs. In each simulation run, a pair of taxis are chosen at random, each of which corresponding to the message source and destination. By following the vehicle contacts in the trace data, the messages are disseminated through a series of taxi or bus contacts towards the destination, according to different strategies as described above. Our goal is to achieve an end-to-end propagation delay less than 8 hours, the number of hours during which people’s daily work is scheduled.

B. Performance Evaluation

1) End-to-End Delay vs. Delivery Ratio: These two performance metrics are evaluated from the message point of view, i.e., how long and how likely a particular message can be delivered within the TTL constraint. Figure 6(a) and (b) show the end-to-end delay and delivery ratio given different message TTL values. Although our target TTL is 8 hours, a wider range of 4 to 20 hours is shown for comparison. Because every successful message dissemination has to be accomplished within the constraint of TTL, given a shorter TTL, the end-to-end delay is shorter, but only a smaller percentage of messages can be delivered successfully.

Our schemes that utilize the selected bus backbone reduce the delivery delay significantly, while the delivery ratio with a small TTL is much higher when compared with the single-copy T2T scheme. The improvement can be seen at the target TTL of 8 hours, where the delivery ratio of both T2B2T schemes are higher than 90%. However, for T2T to have a similar performance, the multi-copy strategy has to be used. As shown in Fig. 6(a) and (b), only when source taxis disseminate
multiple copies of the messages whenever they encounter other taxis, the T2T performance metrics in terms of delivery ratio and delay become comparable to those of T2B2T.

Note that when TTL is smaller than 12 hours, corresponding to 7AM to 7PM during a day, both performance metrics in Fig. 6(a) and (b) increase linearly with TTL. When TTL goes up to 16 hours, i.e., also covering the time from evening to midnight, taxis become more active due to people’s leisure activity at night. Therefore, both the contacts among taxis, and the contacts between taxis and buses, become more frequent. As a result, the oracle model is more likely to find a path with lower delay. We can thus observe a slight decrease in the end-to-end delay in Fig. 6(a). Similar irregularity in Fig. 6(b) is also due to this change of activity in vehicles.

2) Network Traffic: This performance metric is from the network point of view. Figure 7(a) shows the network traffic generated by different message dissemination schemes with different TTL, in terms of the total number of message copies in the network. When using the single-copy forwarding among taxis, there is only one message copy in the system at any given time. When the message forwarding from buses to taxis is controlled, the total network traffic is bounded by the total number of buses, which is even lower than the traffic in multi-copy T2T. Figure 7(b) plots the CDF of network traffic when TTL=8 hours, where the controlled T2B2T has a significantly lower network overhead when compared with the multi-copy T2T. From Fig. 6 and Fig. 7, the T2B2T schemes have a better performance than single-copy T2T. Further, if the T2T scheme is tuned to match the performance of T2B2T, the latter has a much lower network overhead.

V. CONCLUSIONS

In this paper, we propose a multi-modal VANET message dissemination scheme that leverages the wide time and space coverage of taxis, and the high route and schedule regularity of buses in a large modern city. The results of the trace-driven simulation show that both the end-to-end delay and delivery ratio could be significantly improved with the assistance of public bus systems, which indicates the great potential of exploring heterogeneous mobility patterns in mobile ad hoc networks. Our future work includes exploring larger sets of GPS data, and employing better propagation control to reduce the network traffic in the bus backbone.

REFERENCES