An evaluation of vehicular networks with real vehicular GPS traces

Yanmin Zhu\textsuperscript{1,2,*}, Chao Chen\textsuperscript{1} and Min Gao\textsuperscript{3}

Abstract

Vehicular networks have attracted increasing attention from both the academy and industry. Applications of vehicular networks require efficient data communications between vehicles, whose performance is concerned with delivery ratio, delivery delay, and routing cost. The most previous work of routing in vehicular networks assumes oversimplified node mobility when evaluating the performance of vehicular networks, e.g., random mobility or artificial movement traces, which fails to reflect the inherent complexity of real vehicular networks. To understand the achievable performance of vehicular networks under real and complex environments, we first comprehensively analyze the affecting factors that may influence the performance of vehicular networks and then introduce four representative routing algorithms of vehicular networks, i.e., Epidemic, AODV, GPSR, and MaxProp. Next, we develop an NS-2 simulation framework incorporating a large dataset of real taxi GPS traces collected from around 2,600 taxis in Shanghai, China. With this framework, we have implemented the four routing protocols. Extensive trace-driven simulations have been performed to explore the achievable performance of real vehicular networks. The impact of the controllable affecting factors is investigated, such as number of nodes, traffic load, packet TTL, transmission range, and propagation model. Simulation results show that a real vehicular network has surprisingly poor data delivery performance under a wide range of network configurations for all the routing protocols. This strongly suggests that the challenging characteristics of vehicular networks, such as unique node mobility, constraints of road topology, need further exploration.

1 Introduction

Vehicular networks have attracted increasing attention from both the academy and industry because of their potential in fostering a wide spectrum of existing applications, such as driving safety, intelligent transport services [1], mobile Internet access, and file sharing [2-4].

Vehicular networks exhibit some similar characteristics in mobile ad hoc networks (MANETs) and delay tolerant networks (DTNs), depending on the density of vehicles. MANETs and DTNs share the advantage of requiring no support of a fixed infrastructure.

When the vehicle density is higher, vehicular networks show stronger similarity with MANETs because of better network connectivity. On the contrary, when the density is lower, vehicular networks show more similarity with DTNs, where the network is subject to more frequent disruption. However, compared with MANETs and DTNs, vehicular networks are more complex because of unique characteristics such as non-uniform node distribution, fast change of topology, and restricted mobility of vehicles.

Applications of vehicular networks require efficient data communications between vehicles. The main performance of data communication in vehicular networks is concerned with delivery ratio, delivery delay, and routing cost. Many routing algorithms have already existed for use in vehicular networks, such as epidemic [5], AODV, GPSR [6], and MaxProp [7].

It is very important to understand the performance of these routing algorithms for vehicular networks. Unfortunately, most previous work of routing in vehicular networks assumes oversimplified node mobility when evaluating the performance of vehicular networks, e.g., random mobility [6,8] or artificial movement traces [9], which have been widely used in MANETs and DTNs.

*Correspondence: yzhu@cs.usst.edu.cn
\textsuperscript{1}Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
\textsuperscript{2}Shanghai Key Lab of Scalable Computing and Systems, Shanghai 200240, China
Full list of author information is available at the end of the article

© 2013 Zhu et al.; licensee Springer. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
They fail to reflect the inherent complexity of real vehicular networks. Several important factors of vehicular networks impact the performance of data delivery in vehicular networks, which should be taken into account, e.g., vehicle mobility, vehicle density, and radio propagation model. Vehicle mobility is restricted by underlying roads and may have a great impact on the resulting network topology and hence the availability of radio links between vehicles. As a result, the routing performance of the vehicular network is affected by vehicle mobility [10,11].

To understand the achievable performance of vehicular networks under real and complex environments, we first comprehensively analyze the factors that may affect the performance of vehicular networks and then introduce four representative routing algorithms of vehicular networks, i.e., Epidemic, AODV, GPSR, and MaxProp. Next, we develop an NS-2 simulation framework [12] incorporating a large dataset of real taxi GPS traces collected from around 2,600 taxis in Shanghai, China. With this framework, we have implemented the four routing protocols.

Extensive trace-driven simulations have been performed to explore the achievable performance of real vehicular networks. The impact of the controllable affecting factors is investigated, such as number of nodes, traffic load, packet TTL, transmission range, and propagation model. Simulation results show that a real vehicular network has surprisingly poor data delivery performance under a wide range of network configurations for all routing protocols. This strongly suggests that the challenging characteristics of vehicular networks need further exploration, including unique node mobility, constraints of road topology.

In this paper, we have made the following technical contributions.

- We have developed an NS-2 framework incorporating a large dataset of real vehicular GPS traces. It enables us to capture both the real characteristics of vehicle mobility and complex features of urban roads.
- We have implemented four representative routing algorithms in the NS-2 framework and explored the performance of vehicular networks using each of the routing algorithms.
- We have performed extensive simulations to reveal the achievable performance of the routing algorithms in a wide range of affecting factors. Particularly, we have also investigated the impact of the radio propagation model that has usually been ignored in simulation study of previous work.

The rest of this paper is organized as follows. Related work is reviewed in Section 2. Section 3 presents the network model and detailed analysis on factors that affect the performance of vehicular networks. Section 4 introduces the four selected routing protocols which are implemented with NS-2. In Section 5, we present our evaluation methodology and simulation setup. Then, evaluation results are shown in Section 6. Finally, we conclude the paper in Section 7.

2 Related work

Recently, much research efforts have been made to each aspect of vehicular works, including data delivery [11,13,14], infrastructure [15], applications [1,16], etc.

With increasingly demand of mobile communication on the wheels, much work has made comprehensive performance comparison among different routing protocols for vehicular networks in city traffic scenarios. In [17], AODV and OLSR have been evaluated and compared for vehicular networks in urban environment under realistic mobility model called vehicular mobility model which is developed in [18]. In [19], a modified ad hoc on-demand distance vector routing (AODV) [8] and a modified greedy perimeter stateless routing (GPSR) [6] are compared under realistic vehicle traces generated by a multi-agent microscopic traffic simulator (MMTS) simulates the traffic over the real road map of a city in Switzerland. With the vehicular traces generated by the traffic flow simulator Videlio for a small area of Berlin, Lochert et al. [20] compared the routing algorithms AODV, dynamic source routing, and a geographic source routing approach. All the traces are used as input to NS-2 for simulating the mobility of vehicles in the network. All the previous work assume that vehicular networks are similar to MANETs, sharing the common characteristics such as good network connectivity and low mobility in small regions.

Real vehicle traces have also been used for evaluation of vehicular networks. A vehicle-based testbed [7] called UMassDieselNet is composed of 30 busses equipped with 802.11b interface and global position service (GPS), which collects the movement traces of the busses that can be used for simulations. Different from generated traces with simple mobility models, real vehicular traces collected are better in reflecting the complications implied by real road traffic, inter-vehicle dependency, and road topology. Vehicular networks behave more like DTNs when the number of vehicles is small, while the vehicles are distributed over a large area. Several routing algorithms have been developed for DTN-like vehicular networks, such as MaxProp [7].

There is a preliminary attempt [21] to understanding the impact of network properties (connectivity, unpredictability, and resource contention) on performance of different communication routing algorithms, AODV, GPSR, MaxProp, and epidemic.
In our work, we try to understand the realistic performance of a vehicular network, with different routing algorithms, and study the impact of important factors such as vehicle density, traffic load, packet TTL, transmission range, and propagation model. To this end, we have collected real vehicular traces from over 4,000 taxis in Shanghai, one of the largest metropolises in China, and use them as the input to our NS-2 based simulation framework for vehicular networks.

3 Affecting factors
In this section, we analyze the important factors that may have great impact on the performance of data delivery of vehicular networks. Among the factors, only a small subset of them can be controlled in simulation study.

- **Node mobility.** Node mobility is one of the most important factor as it decides the dynamic topology of vehicular network and hence the possible communications among vehicles. Simple mobility can be simulated based on random movement patterns, which has widely been used in simulation study. More convincing mobility traces can be generated by traffic generators based on the actual road map of a city.

- **Number of vehicles.** In a given region of constant size, as the number of nodes increases the network connectivity becomes better. When the network connectivity is higher, a vehicular network performs more like a MANET; when the network connectivity is lower, it becomes more like a DTN. As a consequence, as the network connectivity changes, the routing algorithm that best suits the vehicular network varies as well.

- **Node distribution.** In the real world, vehicles are not uniformly distributed in the given region. Hot spots like commercial district and shopping centers can attract more people, which results in higher node densities in these areas. The heterogeneous distributions of vehicles raise a great challenge for design of routing algorithms.

- **Inter-contact time and duration time.** Inter-contact time [10] characterizes the distribution of the interval between two inter-vehicle contacts. It is clear that as the inter-contact time is smaller, the network connectivity is better. The duration time of a contact decides the amount of data can be transmitted within a contact, which is typically small, in the scale of seconds. The previous study [10] shows that the inter-contact time exhibits an exponential decay of in an urban vehicular environment over a large range of timescale.

- **Transmission range.** The transmission range plays a significant role in vehicular networks. It is intuitive that when the transmission range is larger, a vehicle will have more opportunities of communicating with other vehicles. The previous study [22] presents experimental results on the performance of a vehicular network in different scenarios, which show that vehicles equipped with 802.11b can communicate with each other when the distance is up to around 500 m.

- **Radio propagation model.** The radio propagation model in urban areas is complex considering buildings and other obstacles along the roads which have a significant impact on the effective transmission range. NS-2 implements three different radio propagation models [12]: free space model, two-ray ground reflection model, and shadowing model. To study the impact of radio propagation models, we investigate the two-ray ground model and the shadowing model. The parameters of the Shadowing Model are chosen according to city scenarios in [19]: path loss exponent $\beta = 3.5$, and deviation $\sigma_{db} = 6$.

- **Traffic load.** The performance of a vehicular network is also highly affected by the traffic load. When the network traffic load is light, a given packet has a higher probability of being relayed by intermediate nodes; otherwise, it may be queued at intermediate nodes and a long delivery delay can be experienced.

- **Time-to-Live (TTL) of packets.** In practice, each packet is associated with a TTL. Beyond the TTL, a packet is considered useless. As the TTL is larger, a packet has a longer time to reach the destination, and hence, the delivery probability becomes higher.

4 Representative routing algorithms
One of the most important factors impacting the performance of vehicular networks is the routing algorithm adopted by the vehicular network. Routing algorithms have a significant impact on routing path selection of data packets. Given a specific vehicular network, different routing algorithms lead to very different performance. We list the representative routing algorithms for vehicular networks in Table 1.

In a routing algorithm, there are three important processes which may eventually influence the performance of vehicular networks. The first process concerns the establishment of a routing path from the source to the destination. In a dense network like MANET, it is good

| Routing establishment | Forwarding strategy | Maintain or recover strategy |
|-----------------------|---------------------|-----------------------------|
| Proactive             | Routing table based | Beacon                      |
| Reactive              | Greedy              | Greedy forwarding           |
| Per-hop decision      | Semi-controllable   | Store and carry             |

Table 1 Components of routing protocols
to proactively establish a routing before packets are transferred. However, in sparse, delay-tolerant networks like DTN, a connected routing path usually does not exist. It would be a waste to build routing paths with significant communication overhead. The second process is decided by the first process. A table-based forwarding method can be chosen as the forwarding strategy, when a routing path has been established. In this case, packets are forwarded in likely the most efficient way, which leads to a high delivery ratio. As mentioned before, in vehicular networks with sparse density and high mobility, finding a connected routing path is usually difficult, if not impossible. If the first process is absent for routing path establishment, some semi-controllable forwarding and greedy forwarding methods have frequently been used in vehicular networks. Such forwarding methods can increase the delivery performance by consuming more network resources, such as channel occupation and bandwidth. Then third process concerns how routing paths can be maintained and how a routing failure can be coped with. In vehicular networks, different combinations of the processes can obtain varying performance of data transmission.

We have selected four typical routing protocols for evaluation of performance of vehicular networks. We make a brief comparison of four representative protocols in Table 2, and detail discussions on these protocols are as follows.

- **Epidemic.** It is proposed for those partially connected ad hoc networks, which was presented in [5]. With this algorithm, a vehicle randomly selects a packet and forward it to each of the vehicles that it can communicate. As a result, when the network traffic is low, this algorithm can obtain good performance of data transmission; otherwise, the performance suffers because of the high communication overhead.

- **AODV.** It is a representative example of reactive routing protocols. However, AODV may fail frequently due to the sparse feature of a vehicular network. To make a fair comparison with other protocols, we revise it such that a packet may be temporarily stored on nodes if there are no relay nodes.

- **GPSR.** The seminal work [6] proposed the greedy perimeter stateless routing protocol for wireless networks and is frequently used a baseline for vehicular networks. GPSR uses the location of the forwarder and the destination to make forwarding decisions. There are two strategies: greedy forwarding and routing around the perimeter of the region. This semi-controllable per-hop forwarding strategy is also revised with a ‘carry’ option if there are no relays around. Furthermore, the destination can inform the source node with the help of a location service [23]. The advanced greedy forwarding [19] that utilizes the velocity vector information is also implemented in our simulation. The neighbor discovery and location information beacon interval is set to 2 s.

- **MaxProp.** Contact opportunities become more precious for highly sparse vehicular networks. In [7], MaxProp is proposed a multi-copy routing strategy that sorts packets in the buffer and makes forwarding decisions in each contact of two nodes. Neither network topology nor road map is needed to make routing decisions. Contact information is exploited to improve packet delivery. MaxProp uses the delivery likelihood to calculate the path cost to a destination for each packet. Then, it sorts packets by the number of forwarding hops and the path cost. As a result, a packet with a higher probability of reaching its destination is given a high priority.

## 5 Evaluation methodology

This section first introduces the dataset of real vehicular traces, then defines the performance metrics, and finally presents the simulation setup.

### 5.1 Real vehicular GPS traces

In the dataset of real traces, there are over 4,000 taxis in the city of Shanghai, China, which are equipped with GPS receiver. These taxis report their real-time positions to a central server which records all the reported positions with corresponding timestamps and taxi IDs. The whole urban size of Shanghai is around 133 km in length and 69 km in width. We select a subset of the GPS traces covering the whole city of Shanghai for simulations, which are reformatted as input to NS-2. The Nam (Network AniMator) visualized snapshot of 1,000 taxis is shown in Figure 1. As observed in the snapshot, the distribution of nodes is non-uniform.

### Table 2: A comparison of routing protocols

| Routing algorithm | Establishment | Forwarding strategy | Maintain or recover |
|-------------------|---------------|---------------------|---------------------|
| AODV              | Reactive      | Routing table based | Control packets     |
| GPSR              | Per hop       | Semi-controllable(side information) | Greedy forwarding |
| Maxprop           | Per hop       | Semi-controllable (prediction) | Store and carry |
| Epidemic          | Per hop       | Greedy              | Store and carry     |

http://jwcn.eurasipjournals.com/content/2013/1/190

Zhu et al. EURASIP Journal on Wireless Communications and Networking 2013, 2013:190
5.2 Performance metrics
The performance metrics are defined as follows:

- **Packet delivery ratio (PDR).** It is defined as the fraction of packets that are evaluated received by their destination vehicles before the TTL expires.

- **Average delivery delay.** It equals to the average delay of delivery delays of all packets that have been successfully delivered to destinations.

- **Normalized routing load (NRL).** It is defined as the ratio of extra packet transmissions for packets being delivered from the source to the destination. For multi-copy routing strategies, such as Epidemic and MaxProp.

### Table 3 Simulation parameters

| Parameter                  | Configuration          |
|----------------------------|------------------------|
| Network simulator          | NS-2.34                |
| Simulation time            | 14,760 s               |
| Simulation area            | 131 km x 89 km         |
| MAC protocol               | IEEE 802.11 DCF        |
| Traffic application        | CBR                    |
| Pairs of S-D               | 50                     |
| Number of nodes            | 500 to 1,000           |
| Routing implementation     | Epidemic, AODV, GPSR, MaxProp |
| Traffic load (packet size) | 20 to 100 kB           |
| TTL                        | 1,800 to 10,800 s      |
| Transmission range         | 100 to 500 m           |
| Propagation model          | Two-ray ground shadowing |

Figure 1 A snapshot of Nam visualizer: 1,000 taxis distributed in Shanghai, China.

Figure 2 PDR with different nodes number.
MaxProp, packet copies would introduce extra packet transmissions, and thus, their NRLs are larger than 1.

5.3 Simulation setup

This section details the simulation set-up. In simulations, we use the NS-2 simulator of version 2.34. We implement the DTN extension of buffer management for all of the four routing protocols. The buffer size of each vehicle is set to be infinite, and the packets in the buffer queue are sorted by decreasing residual time to expire. The wireless channel interference model is set as the default value in NS-2, and we leave radio contention to be resolved the 802.11 MAC. The bandwidth is set to 2 Mbps. We randomly select 50 pairs of nodes as source and destination. The traffic generated at a node is set to CBR (constant bit rate) and all traffic is generated in the first hour of simulation. The number of total packets originated by the senders is 500. The parameters for simulations are listed in Table 3. In simulation, the default parameters are set as follows: number of nodes = 500; traffic load = 40 KB; TTL of packets = 7,200 s; transmission range = 400 m; propagation model is two-ray ground.

6 Evaluation results

This section presents evaluation results and describes our observations.

6.1 Impact of number of vehicles

The number of nodes increases from 500 to 1,000. We can observe that the PDR of all the routing algorithms is relatively steady in Figure 2. Epidemic performs best, while MaxProp and other two routing algorithms perform worse than Epidemic. For delivery delay in Figure 3,
as the number of nodes increases, the average delivery delay packets using AODV performs better. This is because there is better network connectivity and consequently communication opportunities between nodes enlarge. The interesting point is that the delivery delay using MaxProp first decreases and then increases. The decrease of the delay is due to more opportunities for being relayed to destinations when the number of nodes is larger at the very beginning. However, as the number of nodes continue to increase, the connectivity becomes better. This causes that DTN-lie routing performs worse. This phenomenon will be further confirmed in the study of impact of transmission range.

Another observation is that the delay performance of GPSR, which quickly increases when the number of nodes becomes larger. The previous work [24] has pointed out that non-optimal routing and the routing loop problem are affected by the node degree and location inconsistencies. Especially in our simulation, the vehicular network behaves more complicated in connectivity, location-based routing performs worse because of non-uniform distribution of nodes and highly dynamic network topology.

The NRL of each routing algorithm is shown in Figure 4. As expected, the cost of Epidemic is the largest. MaxProp performs as stable as GPSR, since MaxProp forwards packets with the maximum delivery probability. Also, communication overhead of GPSR is mainly incurred by beacon packets which are determined by the beacon interval.

6.2 Impact of traffic load

We next study the impact of traffic load on the performance of vehicular network and the performance comparison of different routing algorithms is reported in Figures 5, 6, and 7. The PDR decreases and the average
delay increases as the traffic load becomes heavier for most of the routing algorithms, except MaxProp which performs well even when the traffic load is heavy. For NRL in Figure 7, when the traffic load is light, because of the greedy strategy of packet replication, the NRL of Epidemic keeps good. When the traffic is heavy, the NRL of Epidemic decreases. It is because that the network capacity is limited and excessive packets hinders the successful delivery of all packets. For the other two algorithms, as the traffic load increases, the proportion of control packets decreases, and hence, the NRL decreases.

6.3 Impact of TTL
Figures 8, 9, and 10 show the study of the impact of TTL on packet delivery ratio. Except Epidemic, MaxProp performs better than AODV and GPSR. For average packet delay, MaxProp also performs the best, and the other two routing strategies perform worse. In Figure 10, we can see that the cost of Epidemic increases as the TTL increases. In comparison, the rest two routing algorithms perform relatively stable as the TTL increases.

6.4 Impact of transmission range
In Figures 11, 12, and 13, we study the impact of transmission range. We can see that AODV, GPSR, and Epidemic perform better as the transmission range is larger. MaxProp performs inversely, which is reasonable due to the greedy strategy of packet replication that causes significant contentions. The average delay of other three algorithms all decreases when the transmission range increases which makes the network connectivity better. On the contrary, the average delay of packets increases.
for AODV. This is caused by the extremely poor delivery ratio of AODV (see Figure 11). AODV performs poorly with the communication range of 100 m. When network connectivity becomes better with larger communication ranges, the communication between two nodes over a long distance is feasible. Figure 13 shows that the NRL of AODV increases dramatically with the increasing communication range, which is caused by broadcast of control packets (RREQ) [19]. The NRL of MaxProp and GPSR are modest.

6.5 Impact of propagation model
We study the performance under both the two-ray ground model and the shadowing propagation model. As shown in Figure 14, AODV and GPSR under the more realistic propagation model perform worse than under the simple model in terms of PDR. In Figure 15, the delivery delay becomes smaller under the shadowing propagation model. This is a result of less channel contention. Meanwhile, more overhead is incurred when using the shadowing model, as shown in Figure 16. This is because the link connectivity is not decided only by the transmission range but also the propagation model. Thus, routing algorithms perform more poorly, especially for AODV.

6.6 Summary
Although the DTN-like routing algorithm MaxProp performs better than the other two routing algorithms under many configurations, its performance is still not good. A high delivery ratio is achieved by Epidemic, but high communication overhead is incurred. In summary, each of the factors mentioned above should be taken into account when designing routing strategies for vehicular networks.

7 Conclusions
In this paper, we have presented an evaluation of vehicular networks with a large dataset of real vehicular GPS traces collected from over 2,600 taxis in Shanghai, China. We have studied four representative routing algorithms, i.e., Epidemic, AODV, GPSR, and MaxProp. In addition, we have investigated a comprehensive set of affecting factors, including vehicle density, traffic load, TTL, transmission range, and propagation model. We have developed an NS-2 model to enable the use of real vehicular traces for realistic simulation study and to support detail simulation of radio propagation and link layer behaviors such as link-layer retransmissions.

Simulation results show that vehicular networks pose unique challenges to routing algorithms. All routing algorithms perform poor in a wide range of network configurations. This suggests that routing design for vehicular networks need further to take the complexity and uniqueness of vehicular networks into account.

Competing interests
The authors declare that they have no competing interests.

Acknowledgments
This research is supported by NSFC (nos. 61170238, 60903190, 61027009, 60933011, 61202375, 61170237), Shanghai Pu Jiang Talents Program (10PJ1405800), Shanghai Chen Guang Program (10CG11), MIT of China (2009ZX03006-001-01), Doctoral Fund of Ministry of Education of China (20100073120021), National 863 Program (2009AA012201 and 2011AA010500), HP IRP (CW267311), SJTU SMC Project (201120), STCSM (08dz1501600, 12ZR1414900), Singapore NRF (CREATE E2S2), and Program for Changjiang Scholars and Innovative Research Team in Universities of China (IRT1158, PCSIRT).

Author details
1 Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. 2 Shanghai Key Lab of Scalable Computing and Systems, Shanghai 200240, China. 3 HKUST Fok Ying Tung Graduate School, Guangzhou 51070, China.
