Computer Networks 57 (2013) 2549–2563

Contents lists available at SciVerse ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

TMA: Trajectory-based Multi-Anycast forwarding for efficient multicast data delivery in vehicular networks

Jaehoon (Paul) Jeong [a,*], Tian He [b], David H.C. Du [b]

[a] Department of Software, Sungkyunkwan University, Republic of Korea
[b] Department of Computer Science and Engineering, University of Minnesota, USA

A R T I C L E   I N F O

Article history:
Received 25 October 2012
Received in revised form 23 March 2013
Accepted 8 May 2013
Available online 20 May 2013

Keywords:
Vehicular network
Road network
Multicast
Anycast
Data forwarding
I2V

A B S T R A C T

This paper describes Trajectory-based Multi-Anycast forwarding (TMA), tailored and optimized for the efficient multicast data delivery in vehicular networks in terms of transmission cost. To our knowledge, this is the first attempt to investigate the efficient multicast data delivery in vehicle networks, based on the trajectories of vehicles in the multicast group. Due to the privacy concern, we assume only a central server knows the trajectory of each vehicle and the estimated current location of the vehicle. Therefore, after receiving a request of multicast data delivery from a source vehicle, the central server has to figure out how the data has to be delivered to the moving vehicles in the multicast group. For each target vehicle in the multicast group, multiple packet-and-vehicle rendezvous points are computed as a set of relay nodes to temporarily hold the data, considering the vehicle's trajectory. This set of rendezvous points can be considered an Anycast set for the target vehicle. We have formulated the multicast data delivery as the data delivery to the anycast sets of the multicast group vehicles. Through theoretical analysis and extensive simulation, it is shown that our design provides an efficient multicast for moving vehicles under a variety of vehicular traffic conditions.

1. Introduction

Vehicular Ad Hoc Networks (VANETs) have become one of key components in Vehicular Cyber-Physical Systems for Intelligent Transportation Systems (ITSs) [1–5]. This is because VANET can support the *in situ delivery* of data messages for emergency information dissemination (e.g., accidents and driving hazards), real-time traffic estimation for trip planning, and mobile Internet services. Espectially, for the driving safety (e.g., collision warning message delivery), VANET is more prompt and reliable than cellular networks (e.g., 3G and 4G-LTE) having an additional delay due to data relay via base stations. Also, to support various road network services with cellular networks while servicing the data and voice traffic generated by cellular phones and smartphones, the service providers of the cellular networks will have to spend significant expenses for the infrastructure expansion and service maintenance due to those additional road network services. Based on these observations, VANET is considered worthy of specialized wireless networks for road network services.

For a varity of road network services for the driving safety and efficiency, VANET can leverage the wireless communications for up-to-date data sharing among vehicles having common interests, such as the images or video clips of driving hazard spots, congested areas, and street parking lots. This will be realized through (i) the standardization of Dedicated Short Range Communications (DSRC) [6] for vehicular communications, (ii) the popular demand of GPS navigation systems [7] for the efficient driving, and (iii) the participatory sensing through smartphones or computer vision devices for vehicle safety (e.g., Mobileye

* Corresponding author. Tel.: +82 31 299 4957.
E-mail addresses: pauljeong@skku.edu (Jaehoon (Paul) Jeong), tianhe@cs.umn.edu (T. He), du@cs.umn.edu (D.H.C. Du).

1389-1286/$ - see front matter
http://dx.doi.org/10.1016/j.comnet.2013.05.002

[8]). Therefore, with this trend, we can raise one natural research question of how to take advantage of these GPS-guided driving paths (called *vehicle trajectories*) to efficiently share data in vehicular networks in terms of minimal wireless communication cost.

In this paper, we take advantage of *vehicle trajectories* to efficiently dispatch messages or data to a group of vehicles (defined as *multicast group vehicles*) that have common interests (e.g., road conditions along the driving paths and street parking scenes in urban areas). For examples, we envision that (i) intelligent driving guidance and (ii) location-based service are viable applications. First of all, it is assumed that some of vehicles or smartphone users (as participatory sensors) report regularly their sensing information (e.g., road conditions and environments) collected from their various sensors (e.g., accelerometer [9] and mono-camera [8]) to Traffic Control Center (TCC) [10]; note that TCC is a central server maintaining the vehicle trajectories for the location management for the data delivery toward mobile vehicles like Mobile IP. First, for the intelligent driving guidance, when a road segment is congested, TCC is aware of cars that will go through this segment, based on their trajectories. TCC can notify these cars of this congestion along with the video clip, image or statistics of this congested road segment in the multicast data delivery so that they can select another better moving path beforehand. Second, a location-based service is a targeted information sharing through the up-to-date photos including the gas prices of gas stations among vehicles that may go through the nearby region. It is desirable for this information to reach those relevant vehicles earlier for their possible visit in a wireless-network-bandwidth efficient way, such as multicasting. Note that both applications have to be aware of vehicle trajectories and it is overkill to use broadcast radio to target a fixed set of vehicles having common interests instead of the DSRC communications [6]. On the other hand, the current multicast approaches [11,12] for vehicular networks are not fully addressing this important property of vehicle trajectory to support our target applications for the efficient utilization of wireless channel.

Our paper proposes Trajectory-based Multi-Anycast forwarding (TMA), tailored and optimized for the efficient multicast data delivery in vehicular networks in terms of transmission cost (i.e., the number of transmissions). To the best of our knowledge, our TMA is the first attempt to investigate the vehicle trajectory for the efficient multicast data delivery.

For an efficient multicast, we have the following two challenges. The first challenge is how to select packet-and-vehicle rendezvous points for multicasting. With the vehicle travel delay and packet delivery delay distributions, our TMA algorithm determines multiple rendezvous points (a set of relay nodes to temporarily hold data packets) of the destination vehicle and the packet. These rendezvous points are called *target points* in this paper and can be considered an Anycast set for the destination vehicle. Thus, we formulate the multicast data delivery to multiple destination vehicles in the multicast group as to deliver data to any target points in the anycast sets of those destination vehicles.

The second challenge is how to connect these anycast sets by selecting one target point per anycast set (called representative target point) into a multicast tree, guaranteeing a given data delivery ratio. Our TMA algorithm constructs a Delivery-Ratio Constrained Minimum Steiner Tree with the representative target points for a multicast tree with a minimum channel utilization [13]. Once the multicast tree is constructed, a packet with the multicast tree encoded is source-routed to the target points corresponding to the relay nodes that will hold and deliver the packet to the multicast group vehicles.

Our intellectual contributions are as follows:

- **A multicast data delivery architecture in vehicular networks.** The architecture supports a macro-scoped multicast for multicast group vehicles moving on the different road segments of a target road network.
- **An optimal target point selection algorithm for a multicast group.** This algorithm minimizes the number of target points for the multiple destination vehicles in the multicast group, while guaranteeing the user-required data delivery ratio.
- **A multicast tree construction algorithm for a target optimization goal.** With the selected target points, a multicast tree per packet is constructed to minimize the overall multicast delivery cost or delivery delay, considering the mobility of the multicast group vehicles at the packet transmission time.

The rest of this paper is organized as follows: Section 2 summarizes the related work. Section 3 describes the problem formulation. Section 4 explains the packet and vehicle delay models. Section 5 explains our TMA design. Section 6 explains our TMA protocol. Section 7 evaluates our design. Finally, this paper is concluded with future work in Section 8.

2. Related work

Recently, the VANET research has put a lot of attention on the data forwarding for vehicle-to-vehicle or vehicle-to-infrastructure communications [2–5,14,15]. Most of them are focused on the unicast data forwarding in vehicular networks.

Many data forwarding schemes (e.g., *VADD* [2], Delay Bounded Routing [3], and *SADV* [16]) are investigating the layout of road network and vehicular traffic statistics for the multihop Vehicle-to-Infrastructure (V2I) data delivery. *VADD* [2] investigates the data forwarding based on a stochastic model to achieve the *lowest delivery delay* from vehicle to AP. On the other hand, Delay Bounded Routing [3] proposes data forwarding schemes to satisfy the *user-defined delay bound* rather than the *lowest delivery delay*, while minimizing the channel utilization. *SADV* [16] first proposes a forwarding structure leveraging relay nodes for reliable data delivery. *TBD* [17] utilizes vehicle trajectory information along with vehicular traffic statistics for shorter delivery delay and better delivery probability for multihop vehicle-to-infrastructure data delivery. *TSF* [18] first supports the forwarding for multihop Infrastructure-to-Vehicle (I2V) data delivery, based on vehicle trajectory.

For all those existing approaches, they focus on the unicast data forwarding. On the other hand, *TMA* investigates the Vehicle-to-Vehicle (V2V) multicast data delivery, utilizing the trajectories of those multicast group vehicles.

In the past, multicast routing schemes were proposed for Mobile Ad Hoc Networks (MANETs) [19,20]. Jetcheva and Johnson [19] propose Adaptive Demand-Driven Multicast Routing (ADMR). ADMR uses a loosely-structured multicast forwarding tree rooted at a source node that consists of the shortest-delay paths from the source node to the multicast group members. The tree branches are interconnected with mobile nodes and are locally repaired when some of them are broken due to the mobility of intermediate nodes. However, in vehicular networks to allow for the link breakage and delay tolerance due to the high vehicle mobility, the connection-oriented multicast tree of ADMR is not feasible for a large-scale road network.

Lee et al. [20] propose On-Demand Multicast Routing Protocol (ODMRP). ODMRP uses a *multicast mesh* for richer connectivity among multicast group members rather than a *multicast tree*. The broken branches of the multicast mesh are locally repaired in the similar way with ADMR. As a remarkable element in ODMRP, *mobility prediction* to predict link breakage adapts refresh interval for multicast mesh maintenance to network environments, such as moving speed and pattern. ODMRP is suitable for the multicast among vehicles moving over the same road segment, however it is not suitable for the multicast among vehicles multi-intersection away in the road network. Thus, these legacy multicast schemes (e.g., ADMR and ODMRP) can be used for the *micro-scoped multicast* where vehicles are moving on the same road segment. On the other hand, our TMA is targeted for the *macro-scoped multicast* where vehicles are separately moving over multi-hops in terms of intersections in the road network.

For the multicast in vehicular networks, Sebastian et al. [11] propose an efficient multicast dissemination scheme for the driving safety. The proposed scheme constructs a *micro-scoped multicast tree* consisting of vehicles moving on the same road segment, which is called *micro-scoped multicast*. On the other hand, our *TMA* constructs a *macro-scoped multicast tree* consisting of relay nodes at intersections in the target road network, which is called *macro-scoped multicast*. Kihl et al. [12] propose a reliable geographical multicast routing in vehicular ad hoc networks. This multicast routing forms a multicast tree by using a reactive route discovery for the multicast group vehicles. This approach is not viable for a large-scale road network due to the overhead of the control messages for the route discovery for the multicast group vehicles. On the other hand, for the construction of a macro-scoped multicast tree, *TMA* takes advantage of the trajectories of the multicast group vehicles to identify their locations in a road network without any control messages.

For the data sharing utilizing vehicle trajectories, Leontiadis et al. propose a data forwarding scheme extending Acess Point Connectivity through opportunistic routing from APs to destination vehicles [21,22]. Their approach is similar to our *TMA*, but their forwarding is based on geographically greedy forwarding (looking for a next-hop carrier having a trajectory closer to the destination) rather than the statistical forwarding of *TSF* [18], considering both vehicular traffic statistics and the destination vehicle's trajectory. Due to the greedy forwarding, their scheme cannot guarantee the delivery ratio unlike *TSF*. Also, since their approach allows vehicles to share their trajectories among neighboring vehicles in order to compute their own forwarding metric (i.e., expected delivery delay), there is the privacy exposure of the trajectories, which does not exist in our *TMA* because only TCC maintains the trajectories of the vehicles for the location management of destination vehicles involved in data sharing. Therefore, since our *TMA* adopts the statitiscal data forwarding used in *TSF* for the multicast data delivery, it can guarantee the data delivery ratio and also preserve the privacy of vehicle trajectories.

Note that *TBD* [17] and *TSF* [18] are our prior contributions and *TMA* is based on them. Our *TMA* uses the delay models (i.e., link delay, packet delay, and vehicle delay) in *TSF*. The link cost model in *TMA* is based on the forwarding distance model in *TBD*. In *TMA*, the target point selection for a multicast group is based on the target point selection for a destination vehicle in *TSF*. On top of these previous contributions, we design a *multicast data delivery architecture* for the road-network-wide multicast. In the next section, we will formulate our *TMA* with assumptions and main ideas.

3. Problem formulation

In this section, we formulate the multicast in vehicular networks as follows: *Given a road network with infrastructure nodes (i.e., APs and relay nodes), our goal is to deliver packets reliably from source vehicle (or AP) to multicast group vehicles at the required End-to-End data delivery ratio, while minimizing the delivery cost.*

3.1. The description of vehicular infrastructure

We formally describe the vehicular infrastructure as follows:

- Traffic Control Center (TCC) is a trusted entity that maintains vehicle trajectories without exposing the vehicle trajectories to other vehicles for privacy concerns [10]. TCC determines which AP will disseminate the multicast packet for the multicast group vehicles, as shown in Fig. 1. Note that TCC and APs are interconnected with each other through the wired network.
- Access Point (AP) is a gateway integrating the vehicular network and wired network. AP has the DSRC communications, storage, and processing capability to forward multicast packets from TCC to the multicast group vehicles, as shown in Fig. 1. For the cost effectiveness, APs are *sparsely deployed* into the road network and are interconnected with each other through the wired network or wirelessly (as Mesh Network) [5,23]. Each AP installation with power and wired network connectivity can cost as high as US$5,000 [24].
- Relay Node (RN) is a temporary packet holder for the reliable packet forwarding toward an intended packet forwarding path in a target road network. RN has the

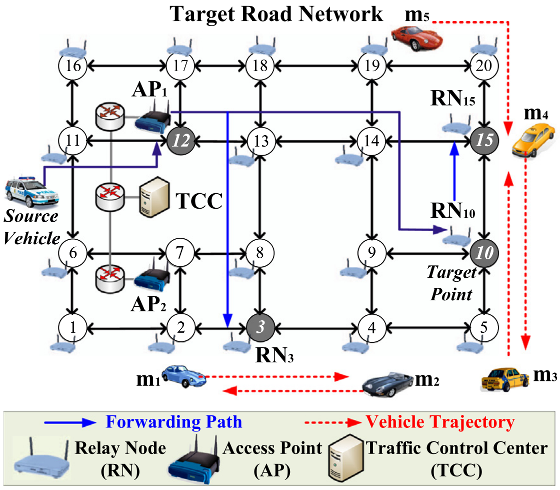


Fig. 1. Multicast forwarding in vehicular network.

DSRC communications, storage, and processing capability, but does not have the wired network connectivity to APs for the cost effectiveness, as shown in Fig. 1. This means that RNs do not have the direct, wired connectivity to either APs or TCC to save deployment cost. Also, it is assumed that RNs are not wirelessly connected to each other. However, in the case where RNs are wirelessly connected, we can regard the road segments among them as wirelessly covered by a Mesh Network consisting of those RNs. With a small number of APs, RNs are used to perform the reliable data delivery from AP to the other RNs corresponding to the target points (i.e., packet destinations) by using intermediate vehicles as packet carriers, moving on road networks. One RN is assumed to be deployed at each intersection for the reliable forwarding, but we can allow some intersections not to have their RNs, discussed in Section 6.2. Of course, RNs can be deployed for the Quality-of-Service (QoS) data delivery in the middle of road segments for a Mesh Network consisting of RNs, but in our TMA, this QoS data delivery is left as future work.

- Vehicles participating in VANET have DSRC device [6]. Nowadays many vehicle vendors, such as GM and Toyota, are planning to release vehicles with DSRC device [25,26]. These vehicles play a role of packet forwarders and packet carriers until they forward packets to a relay node or packet destination vehicle.
- Vehicles, TCC, APs and RNs are installed with GPS-based navigation systems and digital road maps to forward packets to an intended direction [27,28]. Traffic statistics, such as vehicle arrival rate λ and average vehicle speed v per road segment, are available via commercial navigation systems (e.g., Garmin [27]). However, these traffic statistics are used by only TCC to compute a multicast tree.
- Drivers input their travel destination into their GPS-based navigation systems before their travel and so their vehicles can compute their future trajectory based on their current location and their final destination. Multicast-service-participatory vehicles regularly report their trajectory information and their current location to TCC through APs, using the exisiting unicast forwarding scheme, such as SADV [16] and TSF [18].

3.2. Relay-node-assisted forwarding

In this subsection, we justify our vehicular network architecture containing relay nodes. In order to support the *just-in-time* data delivery from AP to destination vehicles, the delivery delay variation in the packet forwarding path should be bounded. Otherwise, the packets will miss the destination vehicles because they may arrive at the target points later than the destination vehicles.

Without relay nodes, the data forwarding schemes based on *stochastic model* (e.g., VADD [2]) cannot be used to reliably deliver packets from AP to mobile destination vehicles. Note that in the stochastic model, each vehicle tries to forward its packets opportunistically to a better neighboring vehicle toward the packet destination, so this packet delivery process is a random walk. However, this stochastic-model-based forwarding has a huge delay variation, so it cannot be used for the multihop infrastructure-to-vehicle data delivery, as shown in [18].

To reduce the delivery delay variation, we deploy relay nodes as packet store-and-forward nodes. In our model, the packet is source-routed via relay nodes at intersections. Our model has a more accurate packet delay model than the stochastic model, so the *just-in-time* delivery can be realized from AP to destination vehicles.

3.3. The concept of Multi-Anycast

We define a new concept of *Multi-Anycast* as follows:

Definition 3.1 (*Multi-Anycast*). Multi-Anycast is the multicast to anycast sets where an anycast set is a set of target points among a multicast group vehicle's future intersections on its vehicle trajectory.

We will explain the concept of Multi-Anycast using Fig. 1. In this figure, *Source Vehicle* sends its data packet to AP (denoted as AP_1) in relay-node-assisted unicast [16,18] where the target point is AP. The AP will multicast the packet to the multicast group vehicles, as shown in Fig. 1. Table 1 shows the trajectories of the multicast group vehicles (m_i for $i = 1, \ldots, 5$) and the corresponding anycast set A_i for m_i in the figure.

AP can send a packet to any target point in anycast set A_i. As a result, the forwarding toward any target point in the anycast set conceptually leads to *Anycast*. For example, as shown in Table 1, vehicle m_1 has its trajectory of

Table 1
Multi-Anycast for multicast group vehicles.

Vehicle	Vehicle trajectory	Anycast set A_i	Target
m_1	$n_2 \to n_3 \to n_4 \to n_5$	$\{n_3, n_4, n_5\}$	n_3
m_2	$n_4 \to n_3 \to n_2 \to n_1$	$\{n_3, n_2, n_1\}$	n_3
m_3	$n_5 \to n_{10} \to n_{15} \to n_{20}$	$\{n_{10}, n_{15}, n_{20}\}$	n_{10}
m_4	$n_{15} \to n_{10} \to n_5 \to n_4$	$\{n_{10}, n_5, n_4\}$	n_{10}
m_5	$n_{19} \to n_{20} \to n_{15}$	$\{n_{20}, n_{15}\}$	n_{15}

$n_2 \to n_3 \to n_4 \to n_5$. For vehicle m_1, the anycast set is $\{n_3, n_4, n_5\}$ that satisfies the delivery ratio α. In this anycast set, any target point can be selected as a packet destination by TCC; note that TCC has the trajectory and location information of each multicast group vehicle. In this figure, TCC selects n_3 as the target point of vehicle m_1.

Our Multi-Anycast problem can be defined as follows: *How to multicast packets to the anycast sets, specifically, to an optimal target point for each anycast set related to each multicast group vehicle at the data delivery ratio* α*, while minimizing the overall multicast delivery cost?*

To answer this problem, in Section 4, we first model the packet delivery delay and the vehicle travel delay used for the estimation of *just-in-time* delivery. Next, in Section 5, with the probability distributions of the packet delivery delay and the vehicle travel delay, we will explain the target point selection and the construction of a multicast tree.

4. Delay models

In this section, we describe three types of delay models for the *just-in-time* delivery, proposed in our TSF scheme [18]: (i) Link delay model, (ii) Packet delay model, and (iii) Vehicle delay model.

4.1. Link delay model

This subsection analyzes the link delay for one road segment with one-way vehicular traffic given the road length (l), the vehicle arrival rate (λ), the vehicle speed (v), and the communication range (R). It is supposed that for packet store-and-forward, one relay node is placed at each endpoint (i.e., intersection) of the road segment. Note that the link delay for a two-way road segment is left as future work.

It is notable that in the VANET scenarios, the carry delay is dominant delay factor because the delay caused by the communication over the air is negligible in comparison to the delay incurred by vehicles carrying the packets. Thus, in our analytical model for the link delay, the carry delay is focused for the sake of clarity, although the small communication delay does exist in our design.

The link delay for one road segment is computed considering the following two cases for the communication range of the relay node at intersection I_i in Fig. 2:

- **Case 1: Immediate Forward**: When the current *packet carrier* n_c arrives at the entrance intersection I_i, there is at least one vehicle (i.e., $k > 0$) moving toward the intended next intersection along the packet's forwarding path. In this case, n_c forwards its packet to the vehicular ad hoc network corresponding to the *Forwarding Distance* l_f in Fig. 2, which consists of k vehicles moving toward the exit intersection I_j of the road segment. First, the packet is forwarded from n_c to n_1 over the *Forwarding Distance* l_f. Next, the packet is carried by n_1 over the *Carry Distance* l_c until n_1 reaches the communication range of the relay node at I_j.
- **Case 2: Wait and Carry**: When the current *packet carrier* n_c arrives at the entrance intersection I_i, there is no vehicle (i.e., $k = 0$) moving toward the intended next intersection along the packet's forwarding path. In this case, n_c forwards its packet to the relay node at the entrance intersection I_i. The relay node holds the packet until another vehicle is moving toward the exit intersection I_j. When a vehicle enters I_i and moves toward I_j, it receives the packet from the relay node I_i and carrys the packet up to the communication range of the relay node at I_j, that is, over the *Carry Distance* of $l - R$.

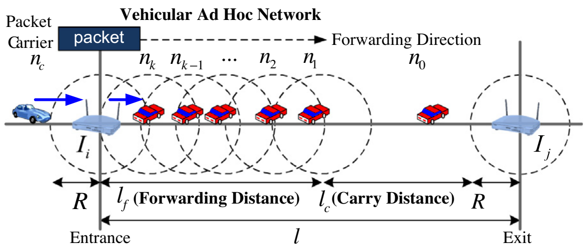


Fig. 2. Link delay modeling for road segment.

Thus, the expectation and variance of the link delay can be computed with the link delays of these two cases as follows:

$$d = \begin{cases} \frac{l - l_f - R}{v} & \text{for case 1: immediate forward,} \\ \frac{1}{\lambda} + \frac{l-R}{v} & \text{for case 2: wait and carry.} \end{cases} \quad (1)$$

$$E[d] = E[\text{link delay}|\text{forward}] \times P[\text{forward}] + E[\text{link delay}|\text{wait}] \times P[\text{wait}]. \quad (2)$$

$$Var[d] = E[d^2] - (E[d])^2. \quad (3)$$

Refer to Appendix B and Appendix C for the detailed derivation of (2) and (3), respectively.

Let $G_r = (V_r, E_r)$ be a road network graph where V_r is the set of intersections and E_r is the matrix of road segments. With the mean $E[d]$ and variance $Var[d]$ of the link delay, we model the link delay d as the Gamma distribution. Note that the Gamma distribution is usually used to model the positive continuous random variable, such as the waiting time and lifetime [29]. Thus, the distribution of the link delay d_i for the edge $e_i \in E_r$ is $d_i \sim \Gamma(\kappa_i, \theta_i)$ such that $E[d_i] = \kappa_i\theta_i$ and $Var[d_i] = \kappa_i\theta_i^2$ for $d_i, \kappa_i, \theta_i > 0$ [29]. Since we have the mean and variance of the link delay, that is, $E[d_i] = \mu_i$ in (2) and $Var[d_i] = \sigma_i^2$ in (3), we can compute the parameters θ_i and κ_i of the Gamma distribution [29].

Note that our design can accommodate an empirical link delay distribution if available through measurement. For this empirical distribution of link delay, adjacent relay nodes can periodically exchange probe packets with each other to obtain link delay samples. Therefore, with the link delay model for a directed edge corresponding to a road segment, we will be able to model the End-to-End packet delay and the vehicle travel delay in the next subsections.

4.2. Packet delay model

In this subsection, we model the End-to-End (E2E) Packet Delay from one position to another position in a given road network. As discussed in Section 4.1, the link delay is modeled as the Gamma distribution of $d_i \sim \Gamma(\kappa_i, \theta_i)$ for edge $e_i \in E_r$ in the road network graph G_r. Given a forwarding path from AP to a target point, we assume that the link delays of edges constructing the path are independent. From this assumption, the mean and variance of the E2E packet delay (*P*) are computed as the sum of the means ($E[P]$) and the sum of the variances ($Var[P]$) of the link delays along the E2E path, respectively. Therefore, the E2E packet delay distribution can be modeled as $P \sim \Gamma(\kappa_p, \theta_p)$ such that $E[P] = \kappa_p\theta_p$ and $Var[P] = \kappa_p\theta_p^2$ for $P, \kappa_p, \theta_p > 0$ [29].

4.3. Vehicle delay model

In this subsection, we model the Vehicle Delay from one position to another position in a given road network in the same way as the Packet Delay Model in Section 4.2. Given the road network graph G_r, the travel time for edge $e_i \in E_r$ is modeled as the Gamma distribution of $t_i \sim \Gamma(\kappa_i, \theta_i)$; note that the *travel time distribution* for each road segment can be obtained through *vehicular traffic measurement* and it is usually considered the *Gamma distribution* by the civil engineering community [30,31]. The parameters κ_i and θ_i of the Gamma distribution are computed with the mean travel time μ_i and the travel time variance σ_i^2 using the relationship among the mean $E[t_i]$, the variance $Var[t_i], \kappa_i$, and θ_i such that $E[t_i] = \kappa_i\theta_i$ and $Var[t_i] = \kappa_i\theta_i^2$ for $t_i, \kappa_i, \theta_i > 0$ [29] in the same way with Link Delay Model in Section 4.1.

Given a vehicle trajectory from the vehicle's current position to a target point, we suppose that the travel times of edges constructing the trajectory are independent. From this assumption, the mean and variance of the E2E vehicle delay (*V*) delay are computed as the sum of the means ($E[V]$) and the sum of the variances ($Var[V]$) of the edge travel delays along the trajectory, respectively. Therefore, the E2E vehicle delay distribution can be modeled as $V \sim \Gamma(\kappa_v, \theta_v)$ such that $E[V] = \kappa_v\theta_v$ and $Var[V] = \kappa_v\theta_v^2$ for $V, \kappa_v, \theta_v > 0$ [29].

So far, we have explained our delay models. In the next section, based on these delay models, we will explain our Multi-Anycast design in detail.

5. Multi-Anycast design

In this section, we explain how to perform Multi-Anycast for multicast group vehicles. We can formulate the optimization of Multi-Anycast data delivery as follows:

Let $G_r = (V_r, E_r)$ be a road network graph where V_r is the set of intersections n_i and E_r is the matrix of road segments e_{ij} whose values are the pairs of physical distance l_{ij} and packet link delay d_{ij} between n_i and n_j. Let M be a set of multicast group vehicles m_i such that $m_i \in M$. Let $V_i = V(n_i)$ be the vehicle travel delay of vehicle m_i from its current position to its target point n_i. Let $P_i = P(n_i)$ be the packet delivery delay from packet source (i.e., AP) to the target point n_i. Let α be a user-defined delivery probability (i.e., delivery ratio), for example, $\alpha = 0.95$. Let A_i be the set of target points (called *anycast set*) for vehicle m_i such that (i) $A_i = \{a_{i1}, a_{i2}, \ldots, a_{is_i}\}$ for $s_i = |A_i|$ and (ii) $Pr[P(a_{ij}) \leqslant V(a_{ij})] \geqslant \alpha$ for $j = 1, \ldots, s_i$. Let a_i^* be an optimal target point in A_i such that the cost from AP to the target point a_i^* is minimum. Let T be a multicast tree for multicast group M. Let $Cost(T)$ be the multicast delivery cost for the tree T; that is, the sum of edge weights in T such that the edge weight is the link channel utilization (i.e., the expected number of transmissions) in the edge, formally defined in Section 5.2.

Our goal is to construct a minimum-cost multicast tree from AP to all multicast group vehicles while guaranteeing a given data delivery ratio α. The following optimization finds an optimal multicast tree T^* to satisfy our goal:

$$T^* \leftarrow \arg \min_{T \subseteq G_r} E[Cost(T)], \tag{4}$$

subject to $Pr[P_i \leqslant V_i] \geqslant \alpha$ for $n_i \in V[T]$ such that $V[T]$ is the vertex set of T and n_i is a target point of vehicle m_i. In (4), an optimal multicast tree T^* is a Delivery-Ratio Constrained Minimum Steiner Tree from packet source *AP* to the target points n_i for all of the multicast group vehicles [13]. Therefore, for a given multicast group M, *Multi-Anycast* can be formally defined as follows:

Multi-Anycast is the packet forwarding scheme from packet source AP to multicast group M with the minimum mulicast delivery cost such that for each anycast set A_i per multicast group vehicle $m_i \in M$, AP multicasts its packet to one target point a_{ij} in the anycast set A_i at the delivery ratio α.

We explain this optimization for the *Multi-Anycast* into the following three steps: First, we explains how to compute an *anycast set of target points A_i* per multicast group vehicle m_i. Second, we describe how to select an *optimal target point a_i^** per anycast set A_i. Last, we explain how to construct a *minimum-cost multicast tree* with the selected target points that satisfies the required data delivery ratio α.

5.1. Step 1: Constructing anycast sets

In this subsection, we explain how to construct an anycast set of target points A_i per multicast group member $m_i \in M$ with the packet delay distribution and vehicle delay distribution. The target point selection is based on the *delivery probability* that the packet will arrive at the target point earlier than the destination vehicle. This delivery probability can be computed with the packet's delivery delay distribution and the destination vehicle's travel delay distribution.

We now explain how to construct anycast set A_i with the trajectory of each multicast group vehicle m_i. Given a required data delivery ratio α (e.g., 0.95), we select target points a_{ij} from the intersections on m_i's trajectory such that $Pr[P(a_{ij}) \leqslant V(a_{ij})] \geqslant \alpha$; note that $Pr[P(a_{ij}) \leqslant V(a_{ij})]$ is the probability that the packet sent by AP will arrive earlier at target point a_{ij} than the destination vehicle m_i. Thus, for each vehicle m_i, we can compute the anycast set A_i, while guaranteeing the required data delivery ratio α as follows:

$$A_i = \{a_{ij} | Pr[P(a_{ij}) \leqslant V(a_{ij})] \geqslant \alpha \text{ for } a_{ij} \in I_i\}, \quad (5)$$

where I_i is a set of intersections a_{ij} on the trajectory of vehicle m_i. For example, as shown in Fig. 3, we assume that there are five vehicles in a multicast group, denoted as m_i for $i = 1, \ldots, 5$. By (5), we can compute an anycast set A_i for each vehicle m_i, as shown in Fig. 3.

Now we explain how to compute the delivery probability introduced by our Trajectory-based Statistical Forwarding called TSF [18]. As a reminder, the packet delay distribution and the vehicle delay distribution can be computed as explained in Section 4.2 and Section 4.3, respectively. The probability distributions of the packet delay P and the vehicle delay V are assumed to be the Gamma distributions such that $P \sim \Gamma(\kappa_p, \theta_p)$ and $V \sim \Gamma(\kappa_v, \theta_v)$ [29]. Assuming that the packet delay distribution and the vehicle delay distribution are independent of each other, the delivery probability $Pr[P_i \leqslant V_i]$ for target point n_i is computed as follows:

$$Pr[P_i \leqslant V_i] = \int_0^{TTL} \int_0^{v} f(p)g(v)dpdv, \quad (6)$$

where $f(p)$ is the probability density function (PDF) of packet delay p and $g(v)$ is the truncated PDF of vehicle delay v with the integration upper bound TTL that is the packet's Time-To-Live (TTL). Note that the delivery probability is computed considering the packet's lifetime TTL; that is, since the packet is discarded after TTL, the probability portion is zero after TTL.

For example, Fig. 4 shows the distribution of packet delay P from access point AP_1 to target point n_{10} (along the forwarding path shown in Fig. 1) and the distribution of vehicle delay V from vehicle m_1's current position n_2 to target point n_{10} (along vehicle m_1's trajectory shown in Fig. 1) where the vehicle m_1's trajectory is $n_2 \to n_3 \to n_4 \to n_5 \to n_{10}$. Note that in Fig. 4, two vertical dotted lines represent the mean of Packet Delay P (i.e., 100 s) and that of Vehicle Delay V (i.e., 150 s), respectively.

Note that by the delivery probability in (6), the target point selection depends on the packet delay model P and the vehicle delay model V that are described in Section 4.

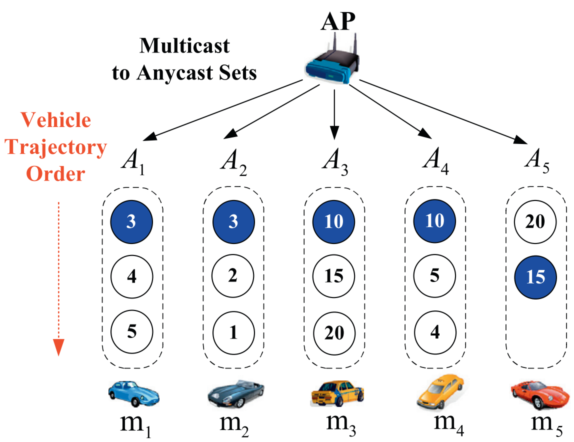


Fig. 3. Anycast sets consisting of target points.

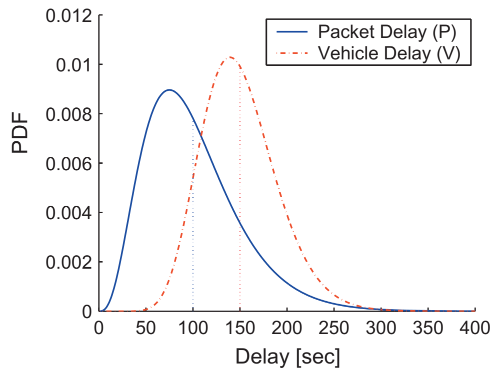


Fig. 4. Packet delay distribution and vehicle delay distribution.

However, our packet delay model and vehicle delay model are not restricted to the Gamma distribution models. Our TMA can accommodate any empirical distributions for both delay models. That is, if more accurate distributions are available, our TMA can use them for the computation of the delivery probability.

5.2. Step 2: Selecting target point points from anycast sets

In this subsection, we explain how to select an optimal target point a_i^* as a *representative target point* for each anycast set A_i per multicast group member $m_i \in M$. We select a target point a_i^* that corresponds to the shortest-path endpoint from source node AP to anycast set A_i in terms of the path cost $Cost(a_{ij})$, which is the sum of edge weights (i.e., link channel utilization values) as follows:

$$a_i^* \leftarrow \arg \min_{a_{ij} \in A_i} Cost(a_{ij}). \quad (7)$$

For example, for five anycast sets in Fig. 3, the representative target points (denoted as colored nodes) are selected such that they are the shortest-path endpoints from AP to A_i for $i = 1, \ldots, 5$. These selected anycast representative target points become packet destination nodes for the multicast tree in the next step, discussed in Section 5.3.

Note that our target point selection algorithm for anycast sets cannot make an optimal set of target points for the overall multicast tree cost. The selection of one target point per anycast set as a destination node for an optimal multicast tree itself is an NP-Complete problem. However, our selection algorithm can make an optimal shortest path tree used as an initial multicast tree in the next step, explained in Section 5.3.

Now, we formally define the *link channel utilization* as *the expected number of transmissions in a road segment*. Fig. 2 shows the forwarding distance l_f of the vehicular ad hoc network in road segment (I_i, I_j), consisting of k vehicles connected by the communication range R. We compute the number of transmissions (denoted as w_{ij}) as follows: $w_{ij} = \lceil E[l_f]/R \rceil$. Refer to Appendix A for the derivation of $E[l_f]$.

5.3. Step 3: Constructing multicast tree for Multi-Anycast

In this subsection, we explain how to build a *Delivery-Ratio Constrained Minimum Steiner Tree* for multicasting data packets to anycast sets where the constraint is the data delivery ratio α. In the previous step, we select the anycast representative target points as the packet destination nodes of the multicast tree. It is known that constructing a Constrained Minimum Steiner Tree itself is an NP-Complete problem. To construct our Constrained Minimum Steiner Tree, we extend Bounded Shortest Multicast Algorithm (BSMA) [13] such that our constraint for the multicast tree is the *data delivery ratio* rather than the *data delivery delay*. Note that since our TMA design is independent of multicast tree algorithms, any efficient multicast tree algorithm can be used to compute a Delivery-Ratio Constrained Minimum Steiner Tree.

As an input for the multicast tree algorithm, we construct the shortest path tree by merging the shortest paths from *AP* to the representative target points a_i^* of anycast sets A_i. Fig. 5a shows the shortest path tree made from five anycast sets A_i for $i = 1, \ldots, 5$. With this initial multicast tree as shown in the figure, the algorithm searches better sub-paths between an arbitrary pair of two multicast nodes (e.g., relay node or destination node) in order to enhance the multicast tree in terms of multicast delivery cost, while satisfying the data delivery ratio α for each target point. Fig. 5b shows a better multicast tree by replacing the path $n_{14} \rightarrow n_{15}$ with the path $n_{14} \rightarrow n_9 \rightarrow n_{10} \rightarrow n_{15}$ to reduce the overall tree cost. Therefore, we can construct a Delivery-Ratio Constrained Minimum Steiner Tree to perform Multi-Anycast to the multicast group.

Note that if the main objective is to provide the shortest delivery delay to the multicast group, our TMA can easily support this objective by using link delay as link cost and selecting a *target point* per multicast group vehicle that provides the shortest delay from AP to the vehicle. *To achieve a shorter multicast delivery delay*, we construct *the shortest path multicast tree* by merging the shortest paths from AP to the selected target points. In this case, we do not use our multicast tree algorithm because this tree already guarantees the shortest data delivery. In the next section, we will explain Multi-Anycast Protocol to deliver data packets to multicast group vehicles using a multicast tree discussed in this section.

6. Multi-Anycast protocol

In this section, we explain our Multi-Anycast forwarding procedure and optimization issues in the multicast forwarding.

6.1. Forwarding procedure

Our Trajectory-based Multi-Anycast (TMA) supports the multicast in the following two phases: (i) V2I unicast from source vehicle to TCC and (ii) I2V multicast from TCC to multicast group vehicles. This TMA's forwarding procedure consists of the following six steps, as shown in Fig. 6.

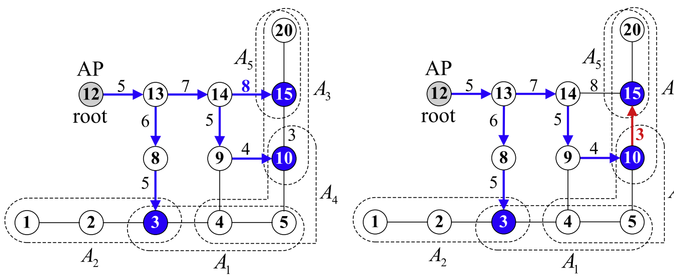
(a) Unicast data forwarding from source vehicle to AP: For multicast data delivery, a source vehicle sends its packet to a nearby AP through the source routing along the shortest path from the source vehicle to the AP, using the exisiting unicast forwarding scheme [16,18].

(b) Unicast data forwarding from AP to TCC: Whenever AP receives a packet for a multicast group, it forwards the packet to TCC having the trajectory information of vehicles. Note that AP and TCC are interconnected through a wired network.

(c) Multicast tree computation per packet at TCC: TCC computes a Delivery-Ratio Constrained Multicast Tree with the vehicle trajectories of the multicast group vehicles through the procedure in Section 5, considering the packet delivery ratio α. TCC then encodes the information of the multicast tree and target points into the packet header using the encoding scheme of multicast forwarding paths, described in [32].

(d) Unicast data forwarding from TCC to AP: TCC forwards the packet with the multicast tree encoded to AP in unicast. This AP is the root of the multicast tree.

(e) Multicast packet forwarding over the multicast tree: AP sends the packet copies to next-hop Relay Nodes (RNs) toward target points along the multicast tree by decoding the packet header. In the same way, intermediate RNs forward the packet copies to next-hop RNs or target points along the multicast tree. Thus, the packet copies



(a) The Shortest Path Tree with Any-cast Representative Target Points

(b) The Delivery-Ratio Constrained Minimum Steiner Tree

Fig. 5. TMA multicast tree construction.

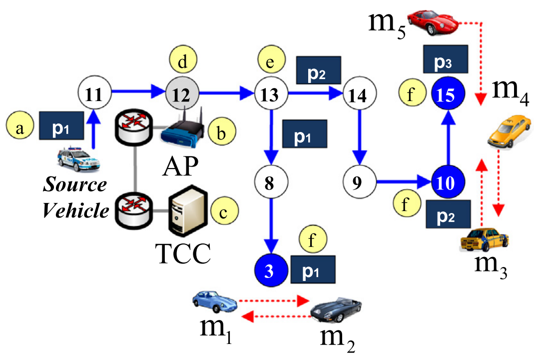


Fig. 6. TMA Multi-Anycast protocol.

can be source-routed from AP to target points (i.e., destination nodes), as shown in Fig. 6.

(f) Store-and-forward at RNs: The packet copies arrive at the RNs corresponding to the destination nodes. The RNs hold packets until the destination vehicles arrive at the intersections having those RNs. When a destination vehicle comes within the communication range of an RN, the packet copies will be forwarded to the destination vehicle. Note that the RN can know when the destination vehicle arrives through RTS/CTS-like handshake because the destination vehicle's identifier (e.g., MAC address) is included in the RTS or CTS frames [33]. This RTS/CTS-like handshake approach may cause packet traffic congestion when too many destination vehicles try to download packets from the same RN. As a result, all of the destination vehicles may not download the packets when they pass through the coverage of the RN. The handling of this congestion issue is discussed in Section 6.2.

To prevent a relay node from holding a packet infinitely, the packet has lifetime as a field in the packet header. In the case where the lifetime of the packet expires, the packet is discarded by the relay node. This lifetime expiration of a packet may happen when the destination vehicle passes through the target point before the packet's arrival at the target point. Note that our target point selection is performed to satisfy the required delivery probability α, as shown in Eq. (5) in Section 5.1. At each target point, a destination vehicle may miss the packet with the probability $1 - \alpha$. To deal with this case, the destination vehicle needs to acknowledge the information of received packets in some way to let the packet source recognize which destination vehicles received the packets. The retransmission mechanism for the reliable multicast delivery is left as future work.

Therefore, our TMA can support the multicast in vehicular networks with the smart combination of V2I and I2V via TCC, having the trajectory information of multicast group vehicles. In the next section, optimization issues for TMA will be discussed.

6.2. TMA optimization issues

We consider the following optimization issues for the practical deployment of TMA systems: (i) TMA forwarding with multiple APs for large-scale road networks, (ii) The scalable TMA systems with multiple TCCs and servers, (iii) The partial deployment of relay nodes, and (iv) The simultaneous packet download by multiple destination vehicles at each target point.

First, with multiple APs, we can support our TMA protocol in a large-scale road network. For each multicast group vehicle, we find an AP among the APs whose delivery cost to the vehicle is minimum and compute the shortest-cost path from the AP to the vehicle. We construct one multicast tree for each AP with the shortest-cost paths originating from the AP. We apply our multicast tree optimization algorithm called BSMA [13] (discussed in Section 5.3) to each multicast tree. With these optimized multicast trees, we can perform multicasting by letting each multicast tree disseminating the packet copies to the multicast group vehicles belonging to the tree at the moment.

Second, in a large-scale road network, one Traffic Control Center (TCC) might not scale up to provide a large number of vehicles with the TMA multicast. TCC can have multiple servers having the replicas of the trajectories and also the large-scale road network can be divided into multiple regions that have their own TCC for the TMA multicast. Each TCC per region performs the TMA multicast in the centralized way with the trajectory information.

Third, the partial deployment of relay nodes allows that some intersections might not have their own relay nodes. In this case, we filter out the edges without Relay Node (RN) from the road network graph. With this filtered graph, we can run our target point selection algorithm in Section 5.1 without any change. Clearly, as the number of relay nodes decreases, the data delivery probability from the AP to the destination vehicle will decrease. Also, it is important to investigate how to deploy the minimum number of relay nodes in order to guarantee the required delivery delay and delivery ratio. This deployment issue is left as future work.

Fourth, an RN for a target point needs to serve multiple vehicles almost simultaneously. For example, many destination vehicles for the same packet can pass through the coverage of the target point almost at the same time. In this case, all of them may not download the packet destined to them from the RN at the target point. This is possible because of the contention based on RTS/CTS handover [33]. To deal with this case, each target point should be selected by considering multiple downloads; that is, each target should allow all the destination vehicles passing through the target point to download the packet. The detailed algorithm for this target point selection is left as future work.

So far, we have explained our Multi-Anycast protocol for the multicast data delivery. Next, we will show the performance of our TMA in a variety of vehicular network settings.

7. Performance evaluation

In this section, we evaluate the performance of *TMA*, performing an optimal target point selection from anycast sets for the multicast tree construction, described in Section 5. The evaluation setting is as follows:

- **Performance Metrics:** We use (i) *Delivery cost*, (ii) *Delivery delay*, and (iii) *Delivery ratio* as metrics. For *Delivery cost* (i.e., the total number of transmissions), we do not count the data delivery cost for the location management for the multicast group vehicles in our performance evaluation. We just count the data delivery cost of a multicasted packet from AP to multicast group vehicles.
- **Baselines:** Our work is the first attempt for the multicast data forwarding based on the vehicle trajectory, so we have no other state-of-the-art schemes for the multicast in large-scaled road networks. To evaluate *TMA*, we compare it with the following two baselines: (i) Random Target Point Selection (*Random*) and (ii) Efficient Network-Wide Flooding (*Flood*). In *Random*, a target point is randomly selected from each vehicle's trajectory. We construct the shortest path tree from AP to the selected target points in terms of delivery delay. In *Flood*, the packet is transmitted toward all the intersections in the target road network. For an efficient flooding, *Flood* multicasts a packet to **all the intersections** in the target road network as target points by *TMA* rather than by a *network-wide broadcasting* for **all the road segments** in the target road network [34].
- **Parameters:** In the performance evaluation, we investigate the impacts of (i) *Vehicular traffic density N*, (ii) *Vehicle speed* μ_v, and (iii) *Vehicle speed deviation* σ_v.

We have built a simulator based on the scheduler provided by SMPL [35] in C with the following settings. A road network with 36 intersections is used in the simulation setting described in Table 2. One Access Point (AP) is deployed in the center of the network and is connected to Traffic Control Center (TCC). Each vehicle's movement pattern is determined by a *Hybrid Mobility model* of City Section Mobility model [36] and Manhattan Mobility model [37] suitable for vehicle mobility in urban areas. From the characteristics of City Section Mobility, the vehicles are randomly placed at one intersection as *start position* among the intersections on the road network and randomly select another intersection as *end position*. The vehicles move according to the roadway from their start position to their end position. Also, the vehicles wait for a random waiting time (e.g., uniformly distributed from 0 to 10 s) at intersections in order to reflect the impact of stop sign or traffic signal. From the characteristics of Manhattan Mobility, as shown in Table 2, the vehicle travel path length l from start position u to end position v is selected from a normal distribution $N(\mu_l, \sigma_l)$ where μ_l is the shortest path distance between these two positions and σ_l determines a random detour distance; this random detour distance reflects that all of the vehicles do not necessarily take the shortest path from their start position and their end position. Once the vehicle arrives at its destination position, it pauses during a random waiting time and randomly selects another destination. Thus, this vehicle travel process is repeated during the simulation time.

On the other hand, among the vehicles, five vehicles are selected as multicast group vehicles, moving around the perimeter areas of the road network according to their vehicle trajectory. Each multicast group vehicle regularly reports its vehicle trajectory and current location to the TCC via AP in the road network. Note that for any road network topology, our TMA can accommodate *any vehicle mobility* because it can accommodate the empirical distributions of packet delivery delay and vehicle travel delay, as discussed in Sections 4 and 5.

The vehicle speed is generated from a normal distribution of $N(\mu_v, \sigma_v)$ [31,38], as shown in Table 2. The *average vehicle speeds* are used in the vehicle speed distribution to generate vehicle speeds for every two directions per two-way road segment; that is, these two average speeds per road segment can be measured from vehicular traffic by dividing the *road segment length* by the *average travel time* over the road segment. For simplicity, we let all of the road segments have the same speed distribution of $N(\mu_v, \sigma_v)$ in the road network for the simulation; note that our design can easily extend this simulation setting to having the variety of vehicle speed distributions for road segments.

As a simple PHY model, it is assumed that all the packets are received correctly if the distance between two nodes (e.g., vehicle, relay node, and AP) is less than the communication range *R*. As a result, all the packets can always be forwarded to the next-hop node if the distance is less than *R*.

During the simulation, following an exponential distribution with a mean of 5 s, packets are dynamically generated by AP in the road network. The total number of generated packets is 1000 and the simulation is continued until all of these packets are either delivered or dropped due to TTL expiration. The system parameters are selected based on a typical DSRC scenario [6]. Unless otherwise specified, the default values in Table 2 are used.

Table 2
Simulation configuration.

Parameter	Description
Road network	The number of intersections is 36. The area of the road map is 2.025 km × 1.8 km (i.e., 1.26 miles × 1.12 miles)
Communication range	*R* = 200 m (i.e., 656 feet)
Number of vehicles (*N*)	The number *N* of vehicles moving within the road network. The default of *N* is 160
Time-To-Live (*TTL*)	The expiration time of a packet. The default *TTL* is the vehicle trajectory's lifetime, that is, the vehicle's travel time for the trajectory, i.e., 630 s
Vehicle speed (v)	$v \sim N(\mu_v, \sigma_v)$ where $\mu_v = \{20, 25, \ldots, 60\}$ MPH and $\sigma_v = \{1, 2, \ldots, 10\}$ MPH. The maximum and minimum speeds are $\mu_v + 3\sigma_v$ and $\mu_v - 3\sigma_v$, respectively. The default of (μ_v, σ_v) is (40,5) MPH.
Vehicle travel path length (l)	Let $d_{u,v}$ be the shortest path distance from start position u to end position v in the road network. $l \sim N(\mu_l, \sigma_l)$ where $\mu_l = d_{u,v}$ km and $\sigma_l = 3$ km (1.86 miles)

7.1. Forwarding behavior comparison

We compare the forwarding behaviors of *TMA*, *Random* and *Flood* with the cumulative distribution function (CDF) of packet delivery cost; note that for *TMA*, the data delivery ratio threshold α is 95%. From Fig. 7, it is very clear that *TMA* has lower delivery cost than *Random*, and also much lower delivery cost than *Flood*. That is, for any given CDF value from the vertical axis in the figure, *TMA* always has lower cost in the horizontal axis than *Random* and *Flood*. For example, *TMA* needs a delivery cost of 16 transmissions for 85% CDF while for the same CDF value, *Random* needs 48 transmissions and *Flood* needs 280 transmissions. In other words, on the delivery cost, *TMA* requires 33.3% transmission of *Random* and 5.7% transmission of *Flood*, respectively. We will show the forwarding performance of these three schemes quantitatively in the following subsections.

7.2. The impact of vehicle number

The number of vehicles in the road network determines the vehicular traffic density in a road network. In this subsection, we intend to study how effectively *TMA* can forward packets from AP toward the multicast group vehicles using their vehicle trajectories. Through our extensive simulations, we observe that under *any vehicular traffic density*, *TMA* significantly outperforms *Random* and *Flood* in terms of the average delivery cost per packet for the multicast group. Fig. 8a shows the packet delivery cost comparison among *TMA*, *Random* and *Flood* with varying the number of vehicles, that is, from 60 to 240. From this figure, *TMA* has lower packet delivery cost than *Random* and *Flood* at all vehicular densities. The observed trend is that the delivery cost in *TMA* is almost stable even though the number of vehicles increases. This is because *TMA* always tries to construct a minimum-cost multicast tree. On the other hand, *Flood* needs higher delivery cost as the vehicular density increases. This is because the higher vehicular density generates more duplicate packets in flooding, leading to the more transmissions. It is observed that *Random* has a constant stable curve similar to *TMA*'s curve due to its randomness in target point selection. For the average transmission number, as shown in Fig. 8a, *Flood* has 14 times more transmissions than *TMA* and *Random* has almost two times more transmissions than *TMA*.

For the delivery delay, as shown in Fig. 8b, as the vehicular density increases, the delivery delay decreases. This is because the more vehicles increase the forwarding probability among vehicles, so this reduces the carry delay, leading to the overall shorter delivery delay. *TMA* has 6 times longer delay than *Flood*, but has only 85% delay of *Random*. From Fig. 8a and b, *TMA* takes 6 times longer delivery delay of *Flood*, but needs only 7% delivery cost of *Flood*. Thus, even though *TMA* sacrifices the delivery delay compared with *Flood*, *TMA* can reduce significantly the delivery cost of *Flood* for multicast data delivery.

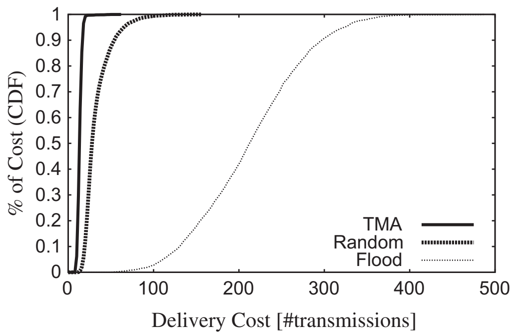


Fig. 7. The CDF of delivery cost.

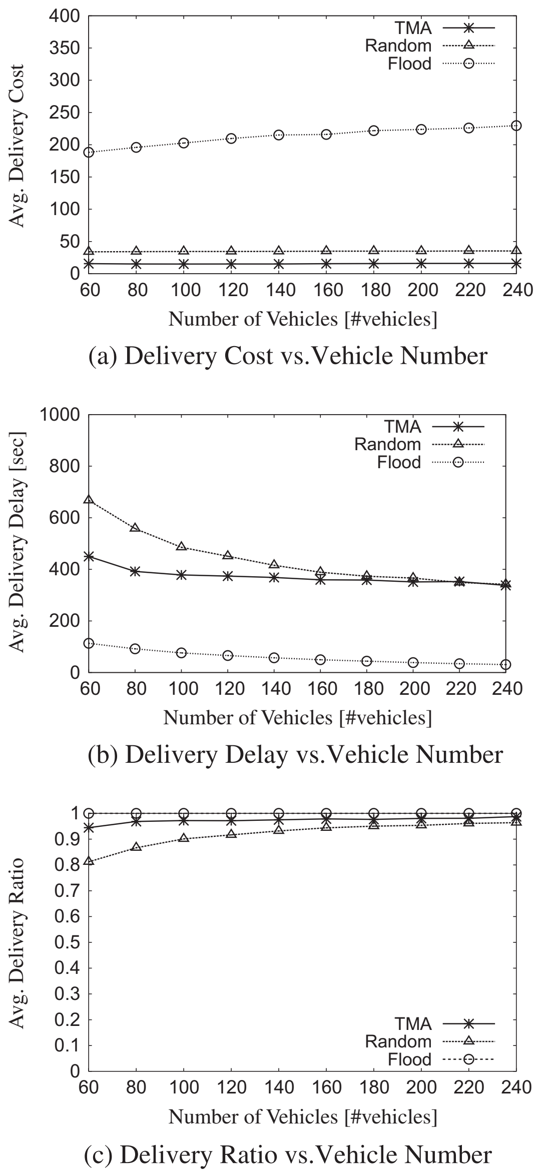


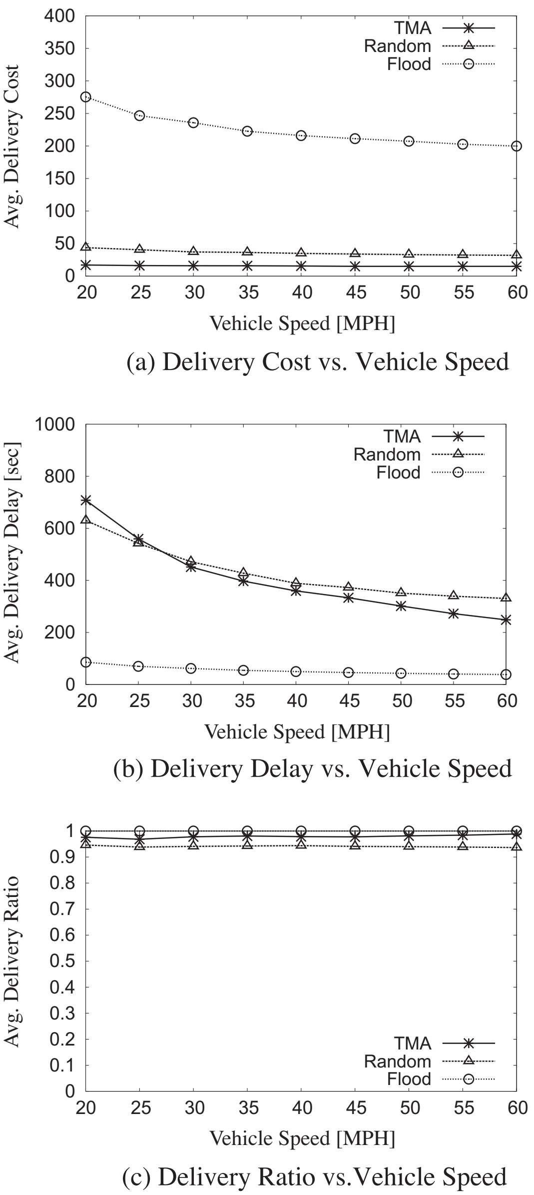
Fig. 8. Impact of vehicle number.

Let us compare the delivery ratios among these three schemes. Fig. 8c shows the delivery ratio for the vehicle number. As discussed for the delivery delay, the higher vehicular density leads to the shorter delivery delay because of the increase of the delivery success probability for the limited packet lifetime (i.e., TTL). Thus, the packet can be delivered to the multicast group vehicles with a higher probability, indicating the high delivery ratio. *Flood* always has 100% delivery ratio regardless of vehicular density, but *TMA* and *Random* have higher delivery ratio as the vehicle number increases. From this figure, it can be seen that *TMA* has more than 95% delivery ratio except low vehicular density (i.e., $N = 60$). As expected, *Random* has lower delivery ratio than *TMA* at all the vehicular densities.

Therefore, through the optimal target point selection for the multicast group vehicles, *TMA* has better performance than *Random* and *Flood* in terms of packet delivery cost. This indicates the importance of an optimal target point selection for the multicast data delivery.

7.3. The impact of vehicle speed

In this subsection, we investigate how the change of mean vehicle speed affects the delivery cost, delivery delay, and delivery ratio. For the delivery cost, as shown in Fig. 9a, *TMA* outperforms *Random* and *Flood*. The higher vehicle speed leads to the lower delivery cost. For *TMA*, it can be seen that the vehicle speed up to 30MPH is helpful to reduce the overall multicast delivery cost, but the higher

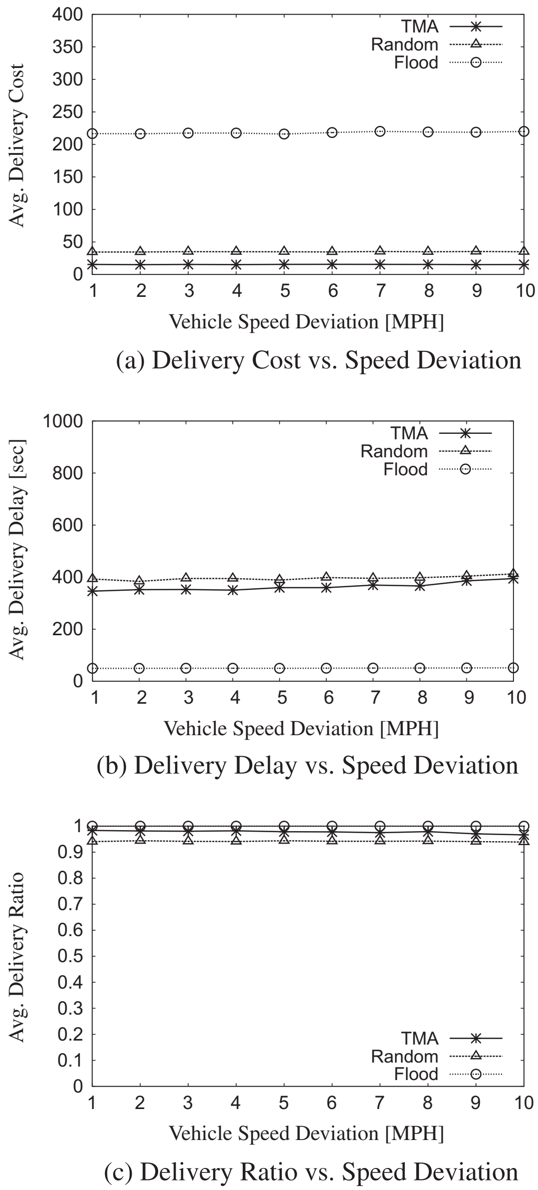


(a) Delivery Cost vs. Vehicle Speed

(b) Delivery Delay vs. Vehicle Speed

(c) Delivery Ratio vs.Vehicle Speed

Fig. 9. Impact of vehicle speed.



(a) Delivery Cost vs. Speed Deviation

(b) Delivery Delay vs. Speed Deviation

(c) Delivery Ratio vs. Speed Deviation

Fig. 10. Impact of vehicle speed deviation.

speed does not contribute much to the delivery cost reduction.

As shown in Fig. 9b, for *TMA*, *Random* and *Flood*, the higher vehicle speed leads to the shorter delivery delay. This is because the high vehicle speed yields high vehicle arrival rate at each road segment, leading to the shorter delivery delay. Note that at low speeds (i.e., 20 and 25MPH), *Random* has shorter delivery delay than *TMA*. This is because *TMA*'s optimization is focused on the delivery cost rather than the delivery delay. For the delivery ratio, as shown in Fig. 9c, *TMA* has better performance than *Random*.

7.4. The impact of vehicle speed deviation

In this subsection, we investigate the impact of vehicle speed deviation on the performance. We found that under a variety of vehicle speed deviations, *TMA* provides a lower delivery cost, a shorter delivery delay, and a higher delivery ratio than *Random*. Also, for the range of vehicle speed deviation, all of three schemes have almost constant curves for three performance parameters. For the packet delivery cost, as shown in Fig. 10a, *TMA*, *Random* and *Flood* have the stable curves over the range of the vehicle speed deviation.

Fig. 10b shows the delivery delay according to the vehicle speed deviation. The delay performance gaps among these three schemes are almost constant at all of the vehicle speed deviations from 1 MPH to 10 MPH. However, for the delivery ratio, as shown in Fig. 10c, *TMA* provides a reliable delivery ratio more than 96%, however *Random* has the worst performance, about 94% delivery ratio in average.

Note that the average delivery delay in our evaluation (e.g., Figs. 8–10) is relatively long compared with the delay in other network domains (e.g., Mobile IP and Mobile Ad Hoc Networks). However, the multimedia data sharing (e.g., image, audio, and video files) for driving safety are tolerant to a little long delay. This is because in many cases, the driving takes more than 10 min (i.e., 600 s). To achieve a shorter delay, we can still use our TMA by selecting target points with the shortest delivery delay and making them into the shortest path multicast tree, as discussed in Section 5.3.

Therefore, it is concluded that *TMA* is a promising approach for the reliable, efficient multicast data delivery in vehicular networks through the *Multi-Anycast* based on the trajectories of multicast group vehicles.

8. Conclusion

In this paper, we propose Trajectory-based Multi-Anycast forwarding (TMA) for multicast data delivery in vehicular networks. Our goal is to provide a reliable, efficient multicast data delivery by minimizing the packet delivery cost (i.e., channel utilization) at the required data delivery ratio. This goal is achieved by computing *packet-and-vehicle-rendezvous-points* (called target points) for the data delivery to multicast group vehicles with the vehicle delay distribution and the packet delay distribution. These distributions can be obtained from the vehicle trajectory and the vehicular traffic statistics. Once optimal target points are determined for the multicast group vehicles, our TMA algorithm constructs a Delivery-Ratio Constrained Minimum Steiner Tree from the AP to the mobile multicast group vehicles. Data packets with the multicast tree encoded are source-routed from AP to the packet destinations along the multicast tree. With GPS-based navigation systems and DSRC communication devices, our TMA shows the effectiveness of vehicle trajectory in the multicast data delivery for the efficient data sharing in vehicular networks. As future work, we will explore the cost-effective deployment of infrastructure nodes (i.e., Access Points) to support Quality-of-Service in large-scale road networks for the given user-required delivery delay and delivery ratio.

Acknowledgments

This research was supported in part by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2012033347). This work was also partly supported by the IT R&D program of MKE/KEIT [10041244, SmartTV 2.0 Software Platform] and by DGIST CPS Global Center. In addition, this research was supported in part by MSI and DTC at the University of Minnesota.

Appendix A. Average forwarding distance

The average forwarding distance $E[l_f]$ can be computed as the expected sum of the inter-distances D_h for $h = 1, \ldots, k$ within the vehicular ad hoc network consisting of k vehicles, as shown in Fig. 2. Suppose that the inter-arrival time T_h is exponentially distributed with arrival rate λ. We have the relationship between the inter-distance D_h and the inter-arrival time T_h such that $D_h = vT_h$. Thus, D_h is also exponentially distributed with λ.

First of all, we define variables for average forwarding distance. Let $a = R/v$; that is, a is the time taken for a vehicle to move out of the communication range R with speed v. Let $C(k)$ be the condition for the connected vehicular ad hoc network consisting of k vehicle inter-arrivals (as shown in Fig. 2) such that $C(k) : T_0 > a$ and $T_h \leqslant a$ for $h = 1, \ldots, k$ where T_0 is the inter-arrival time between n_0 and n_1 in Fig. 2. In the setting of Fig. 2, n_1 is the head vehicle in the vehicular ad hoc network for forwarding the packet of the packet carrier n_c. Note that n_0 arrived at the entrance I_i earlier than n_1 and that n_0 and n_1 are disconnected by the inter-arrival time greater than $a = R/v$. Let $L(k)$ be the length of the connected ad hoc network consisting of k vehicle inter-arrivals. Then, $E[l_f]$ can be derived using the law of total expectation as follows:

$$\begin{aligned} E[l_f] &= E[L] = \sum_{k=1}^{\infty} E[L(k)|C(k)] \times P[C(k)] \\ &= v \times \sum_{k=1}^{\infty} \left(E\left[\sum_{h=1}^{k} T_h | T_0 > a, T_h \leqslant a \text{ for } h = 1, \ldots, k \right] \right. \\ &\quad \left. \times P[T_0 > a, T_h \leqslant a \text{ for } h = 1, \ldots, k] \right) = E[vT_h | vT_h \leqslant R] \\ &\quad \times \frac{P[vT_h \leqslant R]}{P[vT_h > R]} = E[D_h | D_h \leqslant R] \times \frac{P[D_h \leqslant R]}{P[D_h > R]}. \end{aligned} \tag{A.1}$$

For the detailed derivation of $E[l_f]$ along with $L(k)$ and $C(k)$, refer to Appendix in our previous work TBD [17].

Appendix B. Mean link delay

The mean link delay for road segment (I_i, I_j) of length l is computed considering the two cases in Fig. 2: (i) Immediate Forward and (ii) Wait and Carry. Suppose that the vehicles arrive with arrival rate λ. Let $C(k)$ be the condition for the vehicular ad hoc network consisting of k vehicle inter-arrivals. Let $L(k)$ be the length of the connected ad hoc network of k vehicle inter-arrivals. Thus, the mean link delay $E[d]$ is computed by the sum of the conditional expectations for the two cases in (1):

$$\begin{aligned} E[d] &= \left(\sum_{k=1}^{\infty} E\left[\frac{l-R-L(k)}{v} | C(k)\right] \times P[C(k)]\right) \\ &\quad \times P[\text{forward}] + \left(E[\text{waiting time}] + \frac{l-R}{v}\right) \\ &\quad \times P[\text{wait}] \\ &= \frac{l-R-E[l_f]}{v}\beta + \left(\frac{1}{\lambda} + \frac{l-R}{v}\right)(1-\beta), \end{aligned} \quad \text{(B.1)}$$

where $P[\text{forward}] = \beta = 1 - e^{-\frac{\lambda R}{v}}$, $P[\text{wait}] = 1 - \beta = e^{-\frac{\lambda R}{v}}$, and $E[\text{waiting time}] = \frac{1}{\lambda}$. For the detailed derivation of $L(k)$, $C(k)$, and $E[l_f]$, refer to Appendix A.

Appendix C. Variance of link delay

For the variance of link delay, the second moment of link delay $E[d^2]$ is computed as follows:

$$\begin{aligned} E[d^2] &= \left(\sum_{k=1}^{\infty} E\left[\left(\frac{l-R-L(k)}{v}\right)^2 | C(k)\right] \times P[C(k)]\right) \times P[\text{forward}] \\ &\quad + \left(E[\text{waiting time}] + \frac{l-R}{v}\right)^2 \times P[\text{wait}] \\ &= \frac{(l-R)^2 - 2(l-R)E[l_f] + E[l_f^2]}{v^2}\beta + \left(\frac{1}{\lambda} + \frac{l-R}{v}\right)^2 (1-\beta). \end{aligned} \quad \text{(C.1)}$$

Therefore, the link delay variance $Var[d]$ is computed from (B.1) and (C.1) as follows: $Var[d] = E[d^2] - (E[d])^2$.

References

[1] Qing Xu, Raja Sengupta, Daniel Jiang, Design and analysis of highway safety communication protocol in 5.9 GHz dedicated short range communication spectrum, in: VTC, IEEE, April 2003.

[2] Jing Zhao, Guohong Cao, VADD: Vehicle-Assisted Data Delivery in vehicular ad hoc networks, IEEE Transactions on Vehicular Technology 57 (3) (2008) 1910–1922.

[3] Antonios Skordylis, Niki Trigoni, Delay-bounded routing in vehicular ad-hoc networks, in: MOBIHOC, ACM, May 2008.

[4] Jörg Ott, Dirk Kutscher, Drive-thru internet: IEEE 802.11b for "Automobile" users, in: INFOCOM, IEEE, March 2004.

[5] Jakob Eriksson, Hari Balakrishnan, Samuel Madden, Cabernet: vehicular content delivery using WiFi, in: MOBICOM, ACM, September 2008.

[6] Y.L. Morgan, Notes on DSRC & WAVE standards suite: its architecture, design, and characteristics, IEEE Communications Surveys & Tutorials 12 (4) (2010) 504–518.

[7] Hiroki Yomogita, Mobile GPS Accelerates Chip Development. <http://techon.nikkeibp.co.jp/article/HONSHI/20070424/131605/>.

[8] Mobileye, Artificial Vision Technology. <http://www.mobileye.com/>.

[9] TrackTec, Driver Diagnostics. <http://www.track-tec.com/fleetsl/Solutions/DriverDiagnosticsAndSafety.aspx>.

[10] Philadelphia Department of Transportation, Traffic Control Center. <http://philadelphia.pahighways.com/philadelphiatcc.html>.

[11] Alvin Sebastian, Maolin Tang, Yanming Feng, Mark Looi, A Multicast Routing Scheme for Efficient Safety Message Dissemination in VANET, in: WCNC, IEEE, April 2010.

[12] Maria Kihl, Mihail Sichitiu, Ted Ekeroth, Michael Rozenberg, Reliable geographical multicast routing in vehicular ad-hoc networks, LNCS 4517 (2007) 315–325.

[13] Mehrdad Parsa, Qing Zhu, J.J. Garcia-Luna-Aceves, An iterative algorithm for delay-constrained minimum-cost multicasting, IEEE Transactions on Networking 6 (4) (1998) 461–474.

[14] Nilanjan Banerjee, Mark D. Corner, Don Towsley, Brian N. Levine, Relays, base stations, and meshes: enhancing mobile networks with infrastructure, in: MOBICOM, ACM, 2008.

[15] Valery Naumov, Thomas R. Gross, Connectivity-Aware Routing (CAR) in Vehicular Ad Hoc Networks, in: INFOCOM, IEEE, 2007.

[16] Yong Ding, Chen Wang, Li Xiao, A Static-Node Assisted Adaptive Routing Protocol in Vehicular Networks, in: VANET, ACM, 2007.

[17] Jaehoon Jeong, Shuo Guo, Yu Gu, Tian He, David Du, Trajectory-based data forwarding for light-traffic vehicular ad-hoc networks, IEEE Transactions on Parallel and Distributed Systems 22 (5) (2011) 743–757.

[18] Jaehoon Jeong, Shuo Guo, Yu Gu, Tian He, David Du, Trajectory-based statistical forwarding for multihop infrastructure-to-vehicle data delivery, IEEE Transactions on Mobile Computing 11 (10) (2012) 1523–1537.

[19] Jorjeta G. Jetcheva, David B. Johnson, Adaptive demand-driven multicast routing in multi-hop wireless ad hoc networks, in: MOBIHOC, ACM, Long Beach, CA, USA, 2001.

[20] Sung-Ju Lee, William Su, Mario Gerla, On-demand multicast routing protocol in multihop wireless mobile networks, Mobile Networks and Applications 7 (6) (2002).

[21] Ilias Leontiadis, Paolo Costa, Cecilia Mascolo, Extending Access Point Connectivity through Opportunistic Routing in Vehicular Networks, in: INFOCOM, IEEE, San Diego, CA, USA, 2010.

[22] Ilias Leontiadis, Paolo Costa, Cecilia Mascolo, A hybrid approach for content-based publish/subscribe in vehicular networks, Pervasive and Mobile Computing 5 (6) (2009) 697–713.

[23] Vladimir Bychkovsky, Bret Hull, Allen Miu, Hari Balakrishnan, Samuel Madden, A measurement study of vehicular internet access using in situ Wi-Fi networks, in: MOBICOM, ACM, 2006.

[24] Jupiter Research, Municipal Wireless: Partner to Spread Risks and Costs While Maximizing Benefit Opportunities, Technical report, June 2005.

[25] General Motors (GM), Vehicle-to-Vehicle (V2V) Communications. <http://www.gm.com/experience/technology/research/overview/isl/vcim.jsp>.

[26] Toyota Motor Corporation (TMC), TMC Develops Onboard DSRC Unit to Improve Traffic Safety. <http://www2.toyota.co.jp/en/news/09/09/0903.html>.

[27] Garmin Ltd., Garmin Traffic. <http://www8.garmin.com/traffic/>.

[28] Savari Networks, StreetWAVE: Roadside Unit. <http://www.savarinetworks.com/files/StreetWAVE-DS-final.pdf>.

[29] M.H. DeGroot, M.J. Schervish, Probability and Statistics, 3rd ed., Addison-Wesley, 2001.

[30] Abishai Polus, A study of travel time and reliability on arterial routes, Transportation 8 (2) (1979) 141–151.

[31] Donald S. Berry, Daniel M. Belmont, Distribution of vehicle speeds and travel times, in: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, July 1950.

[32] Qing Cao, Tian He, Tarek F. Abdelzaher, uCast: unified connectionless multicast for energy efficient content distribution in sensor networks, IEEE Transactions on Parallel and Distributed Systems 18 (2) (2007) 240–250.

[33] Matthew S. Gast, 802.11 Wireless Networks: The Definitive Guide, 2nd ed., O'Reilly Media, 2005.

[34] Kanitsom Suriyapaibonwattana, Chotipat Pomavalai, An Effective Safety Alert Broadcast Algorithm for VANET, in: ISCIT, IEEE, 2008.

[35] M.H. MacDougall, Simulating Computer Systems: Techniques and Tools, MIT Press, 1987.

[36] Tracy Camp, Jeff Boleng, Vanessa Davies, A survey of mobility models for ad hoc network research, Wireless Communications and Mobility Computing (WCMC): Special Issue on Mobile Ad Hoc Networking: Research, Trends and Applications 2 (2002) 483–502.

[37] Fan Bai, Narayanan Sadagopan, Ahmed Helmy, IMPORTANT: a framework to systematically analyze the impact of mobility on performance of RouTing protocols for Adhoc NeTworks, in: INFOCOM, IEEE, 2003.

[38] Victor Muchuruza, Renatus Mussa, Traffic operation and safety analyses of minimum speed limits on Florida rural interstate highways, in: Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa, USA, August 2005.

Jaehoon (Paul) Jeong is an assistant professor in the Department of Software at Sungkyunkwan University in Korea. He received his Ph.D. degree in the Department of Computer Science and Engineering at the University of Minnesota in 2009. He received the B.S. degree in the Department of Information Engineering at Sungkyunkwan University and the M.S. degree from the School of Computer Science and Engineering at Seoul National University in Korea, in 1999 and 2001, respectively. His research areas are vehicular networks, wireless sensor networks, and mobile ad hoc networks. His two data forwarding schemes (called TBD and TSF) for vehicular networks were selected as spotlight papers in IEEE Transactions on Parallel and Distributed Systems in 2011 and in IEEE Transactions on Mobile Computing in 2012, respectively. He is a member of ACM, IEEE and the IEEE Computer Society.

Tian He is currently an associate professor in the Department of Computer Science and Engineering at the University of Minnesota – Twin Cities. He received the Ph.D. degree under Professor John A. Stankovic from the University of Virginia, Virginia in 2004. He is the author and co-author of over 90 papers in premier sensor network journals and conferences with over 4000 citations. His publications have been selected as graduate-level course materials by over 50 universities in the United States and other countries.

David H.C. Du is currently the Qwest chair professor in the Department of Computer Science and Engineering at the University of Minnesota – Twin Cities. He received the B.S. degree in mathematics from National Tsing-Hua University, Taiwain, ROC in 1974, and the M.S. and Ph.D degrees in computer science from the University of Washington, Seattle, in 1980 and 1981, respectively. His research interests include cyber security, sensor networks, multimedia computing, storage systems, and high-speed networking. He is a fellow of the IEEE.