Delay Analysis and Routing for Two-Dimensional VANETs Using Carry-and-Forward Mechanism

Jianping He, *Member, IEEE*, Lin Cai, *Senior Member, IEEE*,
Jianping Pan, *Senior Member, IEEE*, and Peng Cheng, *Member, IEEE*

Abstract—For disconnected Vehicular Ad hoc NETworks (VANETs), the carry-and-forward mechanism is promising to ensure the delivery success ratio at the cost of a longer delay, as the vehicle travel speed is much lower than the wireless signal propagation speed. Estimating delay is critical to select the paths with low delay, and is also challenging given the random topology and high mobility, and the difficulty to let the message propagate along the selected path. In this paper, we first propose a simple yet effective propagation strategy considering bidirectional vehicle traffic for two-dimensional VANETs, so the opposite-direction vehicles can be used to accelerate the message propagation and the message can largely follow the selected path. Focusing on the propagation delay, an analytical framework is developed to quantify the expected path delay. Using the analytical model, a source node can apply the shortest-path algorithm to select the path with the lowest expected delay. Performance evaluation by simulation show that, when the vehicle density is uneven but known, the proposed Minimum Delay Routing Algorithm can achieve a substantial reduction in delay compared with the geocast-routing approach, and its performance is close to the flooding-based Epidemic algorithm, while our solution maintains only a single copy of the message.

Index Terms—Mobile networks, VANETs, message dissemination, propagation delay, routing

1 Introduction

Equipped with on-board units (OBU), future vehicles can communicate through wireless links. We are particularly interested in vehicular ad hoc networks (VANETs) which can provide low-cost communication services, thank to the low-cost short-range communications and the minimum infrastructure requirement. Although using cellular networks or other wireless infrastructures can provide networking services to vehicles, they often result in a higher spectrum usage fee, or a much more expensive establishing and maintaining cost. In addition, in the areas difficult or too costly to maintain highly reliable infrastructures, or in the situations of natural or man-made disasters when the communication services are most needed while the infrastructures are inaccessible or overloaded, the VANET solution can be a viable option. Thus, we focus our attention on the challenging problem in VANET, i.e., when the vehicle density is low or when the market penetration of OBU is low, network connectivity cannot be guaranteed so that many traditional routing approaches cannot be applied.

- *J. He is with the Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, P. R. China, and the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8P 5C2, Canada. E-mail: jianpinghe.zju@gmail.com.*
- *L. Cai and J. Pan are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8P 5C2, Canada. E-mail: cai@ece.uvic.ca, pan@uvic.ca.*
- *P. Cheng is with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China. E-mail: pcheng@iipc.zju.edu.cn.*

Manuscript received 29 Jan. 2016; revised 1 July 2016; accepted 30 Aug. 2016. Date of publication 9 Sept. 2016; date of current version 1 June 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2016.2607748

Previous research has revealed that for disconnected VANETs, it is promising to apply the "carry-and-forward" mechanism to ensure the delivery success ratio at the cost of a higher delay [1], [2], [3], [4], i.e., when the vehicles are connected as a cluster (vehicles within a cluster can communicate with each other through a multi-hop relay path), messages can be fast relayed within the cluster using wireless communications; when the network is disconnected, messages can be stored and carried by vehicles, and be forwarded to next-hop carriers or destinations later. In this paradigm, two key issues are a) how to find an optimal path from the source to the destination, and b) how the current carrier identifies the suitable next-hop carrier to forward the message along the selected path.

For routing, when there is no guarantee whether a connected path exists, an ideal path depending on carry-and-forward transmissions should ensure successful delivery with the lowest delay: first, with less delay, more time-sensitive applications can be supported; second, a path with less delay means that the number of messages to be carried/stored in the network is smaller, so the storage cost and/or the packet loss rate due to buffer overflow are reduced, and the network capacity is larger.

To find the low latency path, assuming that the communication delay is negligible compared with the message propagation delay when it is carried by a vehicle, intuitively, a path associated with a higher vehicle density will have more chances to fast-forward the message within clusters to reduce the delay. However, roads with high traffic densities often result in a lower vehicle speed (according to the traffic flow theory) and a path with a high vehicle density may be physically longer; both will increase travel time and propagation delay. Thus, how to estimate path delay is essential, which is a challenging, open issue, given the

random topology and high mobility in VANET. Forwarding further complicates the problem, as a message carrier may have its own destination different from the message destination and it is difficult to maintain and store the contact history of all vehicles to identify the ideal next-hop carrier. Even worse, the message may be carried to roads not belonging the selected path. Thus, the random forwarding process introduces more uncertainty in propagation delay.

In the literature, delay performance for emergency message broadcast using both cluster-based fast forwarding and carry-and-forward with vehicles of both directions has been investigated, mostly in one-dimensional high-way scenarios. For instance, [2], [3], [5] considered broadcast applications, where broadcast message propagation direction and the vehicle moving direction are opposite. It is different from our unicast scenario where the message most likely propagates in the same direction as its carrier. In [6], important properties of propagation delay in bidirectional vehicular delay-tolerant networks were revealed, while the delay statistics are still unknown. Furthermore, a large-scale urban VANET is two-dimensional, so the delay components include both the propagation delay along road segments (between two intersections) and the delay in finding the next carrier crossing an intersection towards the desired direction. How to estimate these delay components in two-dimensional VANET to select the path with the minimum expected delay is an open issue.

To fill the gap, this work has the following main contributions. We first propose a simple yet effective propagation strategy for two-dimensional VANETs, so a message can be delivered following the desired path. More importantly, based on the stochastic vehicle traffic model and the propagation strategy proposed, an analytical framework is established to quantify the message propagation delay along the road segments and crossing the intersections. The analysis reveals the relationships among delay components, the vehicle densities and speeds of different directions. Analytical results have been validated by simulation, and they provide important insights and guidance for the routing design. Third, based on the analytical model, a source node can apply the classic shortest path algorithm to obtain the optimal path with the lowest expected delay. We further compare the performance of the proposed Minimum Delay Routing Algorithm (MDRA) with the flooding-based Epidemic algorithm and the popular geocast-routing algorithm. By flooding the message to the whole network, the Epidemic algorithm can achieve the optimal performance in terms of minimizing the delay. The results demonstrate that, when the vehicle density is uneven but known, using the proposed MDRA algorithm can achieve a substantial reduction of path delay when compared with the geocast-routing approach, and its delay performance is close to the flooding-based Epidemic algorithm, while our solution only maintains a single copy of the message in the network.

The rest of this paper is organized as follows. Section 2 discusses the related work. In Section 3, we introduce the system model and the proposed propagation strategy. Section 4 provides some preliminaries. Section 5 gives the delay estimation and the MDRA algorithm design. Performance evaluations by simulation are presented in Section 6, where the impact of the assumptions and how to relax them are discussed, followed by the concluding remarks and future research issues in Section 7.

2 Related Work

Considering its unique feature, many routing solutions have been proposed for VANET, and we only list a small subset of the most relevant ones here. Interested readers are referred to a recent survey [7]. For VANETs with high vehicle densities, it is critical to find the connected paths with low congestion and delay [8]. For sparse VANETs, routing protocols often need to consider the tradeoff between reliability and delay [9]. The carry-and-forward mechanism is promising for sparse VANETs and delay minimization is often the main objective of routing protocols in this paradigm [10], [11]. These works relied on rough estimations of the delay for routing decisions.

Great efforts have been devoted to analyzing the delay performance in VANETs. Ref. [2] analyzed the time to propagate a packet between two disconnected vehicles. Refs. [16] and [17] investigated the expected delivery delay from a vehicle to an Internet access point, and the multi-hop packet delivery delay in a low density VANET, respectively, both considering one-way traffic. In [9], [14], a one-dimensional, two-lane road was considered, and an analytical model was developed to calculate the message delay distribution. Ref. [15] proposed several information release mechanisms to minimize the information delivery delay in the context of an intermittent roadside network, while the intersection and bidirectional roads are not considered. Refs. [12], [13] investigated the delivery delay given the delivery distance and vehicle density, where a one-dimensional road with bidirectional traffic is considered. Ref. [18] provided an accurate analysis framework to estimate the information propagation speed considering one-dimensional road and bidirectional traffic scenario. The previous delay analysis works are mostly confined to one-dimensional roads only, and an accurate analysis on delay statistics for two-dimensional VANET remains unsolved. Modeling and analysis of carry-and-forward delay in VANETs is complicated due to the intermittent connectivity, temporal-dynamic motions, and random forwarding process. As summarized in [7], how to refine the existing models or to design new ones for accurate delay estimation is an open issue, which inspires this work.

3 System Model and Propagation Strategy

3.1 Scenario and Assumptions

Consider a data dissemination problem in VANETs (e.g., the advertisement of a hotel being delivered to the vehicles near the airport), where the objective is to select the best path to forward a message from the source region to the destination region (we consider static destination locations) through vehicles. The vehicle carrying the message will forward it to an appropriate next-hop vehicle based on the selected path and the propagation strategy until the message reaches the destination.

We have a few assumptions. Vehicles can discover others within their communication range through periodic beacon messages. The message delivery information, e.g., the locations of the source and destination, message expiration time, etc., is specified by the source and placed in the packet

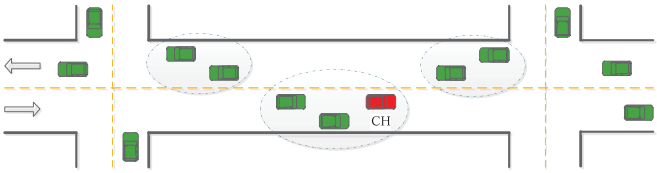


Fig. 1. Example of the vehicle communication model.

header. A vehicle knows its location through GPS or other localization services, and each vehicle encloses its location, speed, and direction information in its periodic beacon messages [10]. We assume that vehicles are equipped with preloaded digital map applications (e.g., MapMechanics [19]), which provide the street-level map and near real-time traffic statistics including vehicle density and speed of different road segments.

3.2 Vehicle Traffic Model and Communication Model

We first introduce the vehicle traffic model. We abstract the VANET to a directed graph $\mathcal{G} = (\mathcal{V}, E)$, where $\mathcal{V}$ is the set of nodes denoting the road intersections, and E is the set of links (edges) denoting road segments between two intersections. Similar to [10], [20], [21], we assume that the vehicles traveling from one intersection to the other follows a Poisson process, with the average arrival rate (per second) $\lambda_{ij} \geq 0$ for $\langle i, j\rangle \in E$, where λ_{ij} can be estimated from the traces in history [20], [22] or given by the map app. Thus, the inter-vehicle arrival times are i.i.d. exponentially distributed. We will investigate the impact of non-Poisson vehicle traffic in Section 6. We assume vehicles in the same direction of a road segment have a constant average speed, denoted by v. (It is straight-forward to plug in the speed in different road segment and direction to the equations with v, so we drop the subscript of v for notation simplicity.) Define the number of vehicles per meter as the vehicle density and assume that the vehicular density along a road segment is known. For vehicles traveling in the direction of $\overrightarrow{ij}$, the corresponding vehicle density is λ_{ij}/v. This traffic model has been used to describe the characteristics of urban traffic scenarios [20], [23]. We also simplify the problem by ignoring the traffic control at intersections, and the impact of this assumption will be discussed in Section 6.

Two vehicles can communicate with each other through a short-range wireless channel, and the communication range is denoted by R. A cluster is defined as the maximal set of vehicles moving in the same direction of a road segment in which every pair is connected by at least one multihop path, where the foremost vehicle is the cluster head.[1] The cluster head, denote by CH, will carry the data when the data is forwarded to the cluster. As shown in Fig. 1, the vehicles in the same circle belongs to the same cluster and the foremost vehicle (e.g., the red one in the middle circle) is the CH. Cluster size is defined as the total length of a cluster. We assume that the vehicles in the same direction and within each other's communication range will self-organize to become a cluster, where the opposite-direction vehicles will not be included in the same cluster. The cluster information is shared by all vehicles in the cluster. We further assume that all vehicles in the same direction of a road segment have the same speed so the clusters remain unchanged until they reach an intersection. (How to maintain or reorganize the cluster considering the network dynamics has been studied, e.g., [24], which is beyond the scope this paper.) For simplicity, we assume that the message exchange time by wireless communications between vehicles in the same cluster is negligible.

1. The cluster head is not the same as those discussed in [2], [18]. Here, the cluster head is always the foremost vehicle in the cluster.

TABLE 1
Important Notations

Symbol	Definition
R	the vehicle communication range
λ_{ij}	the vehicle arrival rate from intersection i to j
λ_{sn}	the arrival rate from south to north at an intersection
c_s	the length or size of a cluster
v	the vehicle speed
X_{ij}	the link propagation delay of $\langle i, j\rangle$
Z_{wn}	the link transfer delay for west to north
Y_{wn}	the message is successfully forwarded from west to north
$\mathbf{E}$	the expectation of random variables
δ	the minimum required time of two clusters to exchange the message
$f_c(x)$	the probability distribution function of cluster size c_s
$\overrightarrow{ij}$	the direction from i to j

Table 1 summarizes a few important notations in this paper for easy reference.

3.3 Propagation Strategy

The propagation strategy includes two parts: how the message is forwarded along a link, called the link propagation strategy, and how the message arriving at an intersection is forwarded to a vehicle traveling in a given direction following the selected path, called the link transfer strategy.

Link Propagation Strategy. Let all vehicles in the cluster have the knowledge of the cluster size and the travel plan of them. The CH will broadcast beacon messages (which include the cluster size) periodically to probe for communication opportunities with other clusters in the opposite direction. We assume that if the heads of two clusters in the opposite directions are within the range of R, they can communicate and exchange message within a time interval of δ which mainly depends on the time of two clusters in the opposite directions capturing the beacon and exchanging handshake messages to initiate data transfer. The CH also knows whether it is within the communication range of the previous cluster it encountered in the opposite direction. Thus, when a CH connects with another CH in the opposite direction, it can know whether the message can be forwarded to its neighbor cluster in front of it with the help of the opposite cluster.

The link propagation strategy can be illustrated using the following example in simulation where vehicles in both directions were generated using Poisson processes and they took the above link propagation strategy. In Fig. 2, the circles in the bottom lane denote the vehicles moving from west to east, while the squares in the top lane denote the vehicles traveling in the opposite direction. The vehicle communication range R is 200 m. The speed of all vehicles is 20 m/s. The message is initially carried by a vehicle at

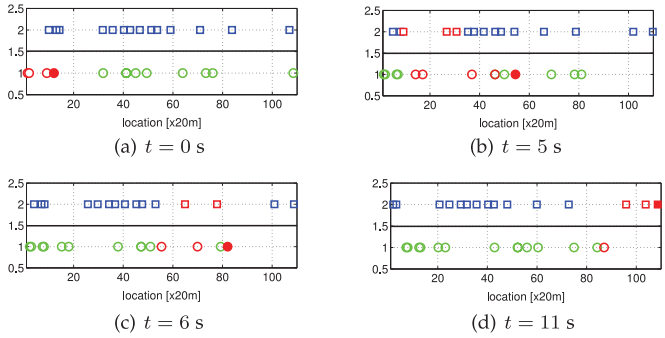


Fig. 2. Example of the link propagation process in a road segment.

location 0 (x-axis), and it needs to be forwarded to location unit 115 (each unit equals 20 m). This message will be forwarded to the CH at location 13 through V2V communications. Thanks to the mobility of vehicles, two neighboring clusters in one direction may be connected through a cluster in the opposite direction. Thus, the message is forwarded to location 55 at time $t = 5$ s and location 84 at time $t = 6$ s by multi-hop communications, and eventually forwarded to location 115 at time $t = 11$ s. In this example, it takes 11 s for the message propagating along the road segment.

Link Transfer Strategy. When a CH carrying the message arrives at a tagged intersection (i.e., the intersection is within the communication range of the CH), it begins to search for the vehicle traveling to the direction following the selected path and then forwards the message to it. Fig. 3 gives an example. The message carrier approaches the intersection at $t = 0$ and needs to find a vehicle traveling to the north. At time $t = 1$, it finds the appropriate vehicle to forward the message.

If failing to find an appropriate vehicle before it leaves the intersection, the carrier will forward the message to the vehicle behind it in the cluster, and this procedure repeats. If the last vehicle in the cluster still cannot find an appropriate next-hop carrier when it passes the intersection, it will keep the message and forward it to the next vehicle encountered that travels in the opposite direction, i.e., traveling back to the tagged intersection. In this way, the message can be carried back to the selected path. This whole process repeats until the message eventually is forwarded to the correct direction at the tagged intersection, so the message can follow the selected path to reach the destination.

4 Preliminaries

In this section, we provide the preliminaries, including the cluster size estimation and forwarding probability, which will be used for the delay estimation in the next section.

4.1 Cluster Size

Let c_s denote the cluster size, which is a random variable. Let λ be the average vehicle density, e.g., $\lambda = \frac{\lambda_{ij}}{v}$ for link $\langle i, j\rangle$. The following theorem is used to estimate the mean and variance of the cluster size, where the approach of the proof has referred to the analysis in [21].

Theorem 4.1. *Suppose that the vehicle density in a path is λ and the inter-vehicle distance follows an i.i.d. distribution with the PDF $f_X(x)$. Then, we have*

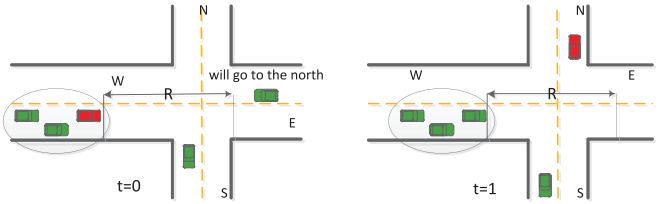


Fig. 3. Example of the link transfer process: car in red is the current message carrier.

$$\mathbf{E}\{c_s\} = \frac{\int_0^R x f_X(x)dx}{\int_R^\infty f_X(x)dx}, \tag{1}$$

and

$$\mathbf{E}\{c_s^2\} = \frac{\int_0^R (2x\mathbf{E}\{c_s\} + x^2) f_X(x)dx}{\int_R^\infty f_X(x)dx}. \tag{2}$$

Proof. Let $X_1, X_2, \ldots, X_n$ be the distances between vehicles 1 and 2, 2 and 3, ..., n and $n+1$, respectively. Clearly $X_1, X_2, \ldots, X_N$ are i.i.d. random variables since the inter-vehicle arrival times are i.i.d. random variables. Then, the average cluster size satisfies

$$\begin{aligned}
\mathbf{E}\{c_s\} =& \sum_{n=1}^{\infty} E\{X_1 + \cdots + X_n | X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall i \le n, X_{n+1} > R\} \\
=& \sum_{n=1}^{\infty} E\{X_1 | X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&+ \sum_{n=2}^{\infty} E\{X_2 + \cdots + X_n | X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall i \le n, X_{n+1} > R\},
\end{aligned}$$

due to the linearity of expectation w.r.t. the sum of random variables. Note that

$$\begin{aligned}
&\sum_{n=1}^{\infty} E\{X_1 | X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall i \le n, X_{n+1} > R\} \\
=& \mathbf{E}\{X_1 | X_1 \le R\} \times \Pr\{X_1 \le R\},
\end{aligned}$$

due to the causality of random variables, and

$$\begin{aligned}
&\sum_{n=2}^{\infty} E\{X_2 + \cdots + X_n | X_i \le R, \forall i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall i \le n, X_{n+1} > R\} \\
=& \sum_{n=2}^{\infty} E\{X_2 + \cdots + X_n | X_i \le R, \forall 2 \le i \le n, X_{n+1} > R\} \\
&\times \Pr\{X_i \le R, \forall 2 \le i \le n, X_{n+1} > R\} \Pr\{X_1 \le R\} \\
=& \mathbf{E}\{c_s\} \Pr\{X_1 \le R\},
\end{aligned}$$

due to the asymptotic property of the conditional expectation of i.i.d. random variables. Thus, one infers from the above three equations that

$$\mathbf{E}\{c_s\} = [\mathbf{E}\{X_1 | X_1 \le R\} + \mathbf{E}\{c_s\}] \times \Pr\{X_1 \le R\}. \tag{3}$$

From the definition of probability, we have $\Pr\{X_1 \le R\} = \int_0^R f_X(x)dx$ and

$$\mathbf{E}\{X_1|X_1 \le R\} = \frac{\int_0^R x f_X(x)dx}{\int_0^R f_X(x)dx}. \quad (4)$$

Substituting the above equation into (3) yields that

$$\mathbf{E}\{c_s\}\left[1 - \int_0^R f_X(x)dx\right] = \int_0^R x f_X(x)dx, \quad (5)$$

which is equivalent to the result shown in (1).

By a similar way, we can derive $\mathbf{E}\{c_s^2\}$. Specifically, $\mathbf{E}\{c_s^2\}$ is derived from solving the following equation,

$$\begin{aligned}\mathbf{E}\{c_s^2\} = (&\mathbf{E}\{c_s^2\} + 2\mathbf{E}\{X_1|X_1 \le R\}\mathbf{E}\{c_s\} \\ &+ \mathbf{E}\{X_1^2|X_1 \le R\}) \times \Pr\{X_1 \le R\},\end{aligned} \quad (6)$$

where $\mathbf{E}\{X_1^2|X_1 \le R\}$ satisfies

$$\mathbf{E}\{X_1^2|X_1 \le R\} = \frac{\int_0^R x^2 f_X(x)dx}{\int_0^R f_X(x)dx}. \quad (7)$$

Substituting (4) and (7) into (6) gives the results that

$$\begin{aligned}\mathbf{E}\{c_s^2\}\left[1 - \int_0^R f_X(x)dx\right] = &2\mathbf{E}\{c_s\}\int_0^R x f_X(x)dx \\ &+ \int_0^R x^2 f_X(x)dx,\end{aligned} \quad (8)$$

which implies that (1) holds. □

The variance of c_s satisfies that $\mathbf{Var}\{c_s\} = \mathbf{E}\{c_s^2\} - \mathbf{E}\{c_s\}^2$. It means that both the mean and variance of c_s can be obtained from Theorem 4.1. Then, we use the following corollary as an example to identify a general approach for the distribution estimation of c_s, which is first given in [21].

Corollary 4.2. *Suppose that the vehicle density in a path is λ and the inter-vehicle distance follows an i.i.d. exponential distribution. Then, the PDF of c_s can be approximated as*

$$f_c(x) = (x^{k-1}e^{-\frac{x}{\theta}})/(\theta^k\Gamma(k)), x > 0, \quad (9)$$

where $k = \frac{\mathbf{E}\{c_s\}^2}{\mathbf{E}\{c_s^2\}-\mathbf{E}\{c_s\}^2}$, $\theta = \frac{\mathbf{E}\{c_s^2\}-\mathbf{E}\{c_s\}^2}{\mathbf{E}\{c_s\}}$, $\Gamma(\cdot)$ is the Gamma function, and $\mathbf{E}\{c_s\}$ and $\mathbf{E}\{c_s^2\}$ are obtained from (1) and (2), respectively.

Proof. Since the inter-vehicle distance follows i.i.d. exponential distribution, we have $f_X(x) = \lambda e^{-\lambda x}$. Substituting the expression of $f_X(x)$ into (1) and (2) can obtain the closed forms of $\mathbf{E}\{c_s\}$ and $\mathbf{E}\{c_s^2\}$, respectively, e.g.,

$$\mathbf{E}\{c_s\} = [1 - (\lambda R + 1)e^{-\lambda R}]/\lambda e^{-\lambda R}.$$

Note that a known Gamma distribution can well approximate and has been widely used to model the sum of exponentially distributed random variables [21]. In this paper, we also use Gamma distribution for approximation of the PDF of c_s. Hence, we have the PDF $f_c(x)$ of c_s satisfies the following form of Gamma distribution,

$$f_c(x) = x^{k-1}\frac{e^{-\frac{x}{\theta}}}{\theta^k\Gamma(k)}, x > 0,$$

where k and θ are two parameters to be determined, and $\Gamma(k)$ is the Gamma function evaluated at k. Note that we have obtained the mean and variance of c_s from Theorem 4.1, which can be used to determine the values of the parameters k and θ in $f_c(x)$. Let $f_c(x)$ have the same mean and variance as those obtained before, we formulate the following equation set

$$\begin{cases} k\theta = \mathbf{E}\{c_s\}, \\ k\theta^2 = \mathbf{E}\{c_s^2\} - \mathbf{E}\{c_s\}^2, \end{cases} \quad (10)$$

and solving this equation set, we obtain that $k = \frac{\mathbf{E}\{c_s\}^2}{\mathbf{E}\{c_s^2\}-\mathbf{E}\{c_s\}^2}$ and $\theta = \frac{\mathbf{E}\{c_s\}}{k} = \frac{\mathbf{E}\{c_s^2\}-\mathbf{E}\{c_s\}^2}{\mathbf{E}\{c_s\}}$. □

Remark 4.3. From the above example, we identified a general approach for the distribution estimation of c_s. Specifically, we can use (1) and (2) in Theorem 4.1 to obtain $\mathbf{E}\{c_s\}$ and $\mathbf{E}\{c_s^2\}$, respectively. Then, use the fact that $\mathbf{Var}\{c_s\} = \mathbf{E}\{c_s^2\} - \mathbf{E}\{c_s\}^2$ to obtain the variance of c_s. Next, solve (10) to obtain the values of the parameters (the parameters of mean and variance) in the given random distribution function of the cluster size. Note that this approach does not depend on a fixed distribution of inter-vehicle distances. Hence, it can be adopted for any kinds of i.i.d. inter-vehicle distance, e.g., following a log-normal distribution [36].

4.2 Link Transfer Probability at Intersections

At an intersection, if the carrier cannot find a suitable vehicle (either within or outside the cluster), the message can be first forwarded to the vehicle behind the carrier within the same cluster who will search for the next carrier, and the procedure can repeat until the last vehicle in the cluster leaves the intersection. Forwarding probability at an intersection is defined as the probability that the message can be successfully forwarded to a vehicle traveling to a given direction before all vehicles in the cluster leave the intersection. This event is denoted by Y and we use subscripts to denote the directions, e.g., Y_{wn} denotes that the cluster approaching the intersection from west can successfully forward the message to a vehicle traveling to north.

Let $R' = R - \delta\Delta v$ be the effective communication range of two clusters being connected, as it takes the time interval of δ for the message to exchange between two clusters, and Δv is the relative speed of the two clusters. The following theorem is given to calculate the probability of Y_{wn}, and its proof is given in Appendix A in the supplementary material, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2016.2607748.

Theorem 4.4. *Suppose that a cluster with size c_s is coming from west carrying the message, and it tries to let a vehicle traveling to north carry the message. The probability of this event to occur can be estimated by*

$$\Pr\{Y_{wn}\} = 1 - (1 - P_w^n)^{|c_s|}e^{-(\lambda_{en}+\lambda_{sn})(c_s+R')/v}, \quad (11)$$

where P_w^n is the probability that a vehicle may turn north from west and $|c_s|$ is the number of vehicles in the cluster approaching the intersection from west.

The above theorem can be applied to both four-way bi-directional intersections and the intersections with more or

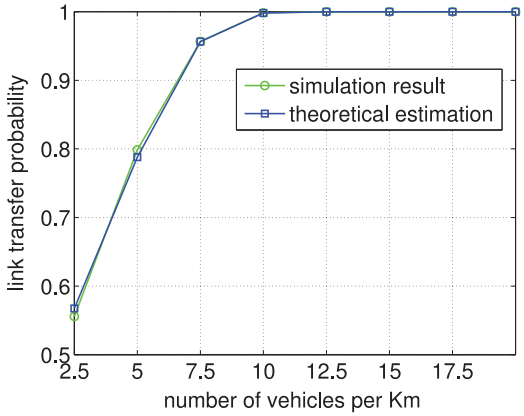
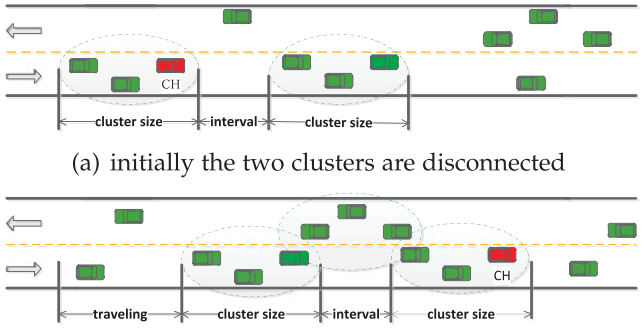


Fig. 4. Validation of the link transfer probability analysis.



(a) initially the two clusters are disconnected

(b) two clusters are connected through the cluster in the opposite direction

Fig. 5. Two clusters getting connected.

fewer roads and with one-way or two-way traffic. For instance, if there are no vehicles or no road from one direction to the other direction, e.g., no traffic from east to north, we set $\lambda_{en} = 0$. P_w^n is given from the traffic statistics. By changing the direction parameters, the forwarding probability of Y between any two directions can be obtained, e.g.,

$$\Pr\{Y_{sn}\} = 1 - (1 - P_s^n)^{|c_s|} e^{-(\lambda_{wn}+\lambda_{en})(c_s+R')/v}. \quad (12)$$

We use simulation to verify Theorem 4.4. We set the vehicle density from each direction to a four-way intersection to be a value varying from 2.5 to 20 (vehicles per km), $R = 200$ m, and $v = 20$ m/s. At the intersection, each vehicle goes straight or turns left or right with equal probability (1/3). (We ignore traffic control at the intersection to simplify the problem, which will be discussed in Section 6.3.) We obtain the successful forwarding ratio from 1,000 simulation runs where vehicles entered the intersection from four directions following Poisson processes. To obtain the analytical results, we set $c_s = \mathbf{E}\{c_s\}$ and $|c_s| = \lceil \mathbf{E}\{c_s\} * \lambda \rceil$, and then use Theorem 4.4 to calculate $\Pr\{Y_{wn}\}$. As shown in Fig. 4, analytical results match well with the simulated ones.

5 Delay Estimation and Minimum Delay Routing

For VANETs relying on carry-and-forward transmissions, minimizing the path delay without sacrificing the reliability is the main goal of the routing algorithm design. To solve this problem, the key issue is to estimate the path delay accurately. Here, we focus on the link propagation delays and the link transfer delays, and ignore the wireless communication delay and losses within the cluster. The link propagation delay is defined as the delay for a message passing through a road segment (between two neighbor intersections), and the link transfer delay is defined as the delay for the message to cross an intersection to the next road segment. In this section, we first investigate the link propagation delay and link transfer delay. Then we estimate the path delay and use it to select the path with the minimum expected delay.

5.1 Link Propagation Delay

Link Propagation Delay Estimation. Link propagation delay mainly depends on the time for a message to move from one CH to the other cluster head or to the end of the link. Define X_c^t the time needed for a message being held by a CH until it reaches the head of its front neighbor cluster with the help of the cluster in the opposite direction. Fig. 5 gives an example, where (a) the message is initially carried by the head of the left cluster and (b) it is then forwarded to the head of the front cluster when the two clusters are connected by the reverse direction vehicles. The message propagation process (moving from one CH to the next) along a link (road segment) is thus a Markov renewal process. X_c^t and the physical distance that the message travels through during X_c^t can be used to estimate the message propagation speed and link delay, so we derive their expected values first.

We first consider the probability of two clusters being connected by vehicles in the opposite direction. Let $x_1 (> R)$ be the distance between two neighbor clusters in the direction $\overrightarrow{ij}$. Recall that the expected cluster size can be obtained from (1) and the distribution of cluster size $f_c(x)$ can be obtained from (9) by setting $\lambda = \lambda_{ji}/v$ for $x \geq 0$. We consider two cases as follows.

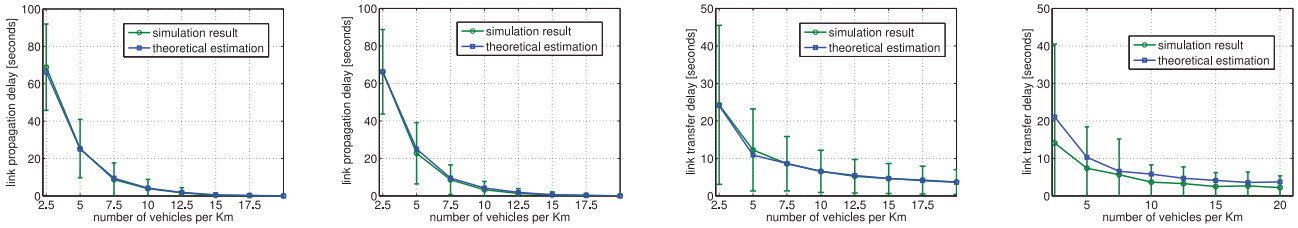
Case 1. $R \leq x_1 \leq 2R'$. If there exist vehicles traveling in the opposite direction in the communication range of both of these two clusters, then $X_c^t = 0$; otherwise, these two clusters can be connected as long as the CH currently holding the message connects to a vehicle from the opposite direction. In this case, the probability of the two clusters being connected before a given time t is

$$\Pr[X_c \leq t | R \leq x_1 \leq 2R'] = P_1(x_1)(1 - e^{-\lambda_{ji} t/2}), \quad (13)$$

where the coefficient $1/2$ is due to the relative speed between vehicles traveling in the opposite directions is twice of v, and $P_1(x_1) = e^{-\lambda_{ji}\frac{2R-x_1}{v}}$ is the probability that the two clusters cannot be connected directly.

Case 2. $x_1 \geq 2R'$. If there is a cluster with $c_s \geq x_1 - 2R'$ in the opposite direction communicating with these two clusters, then $X_c^t = 0$; otherwise, they are connected only when the first cluster meets a vehicle coming from the opposite direction which belongs to a cluster with $c_s \geq x_1 - 2R'$. Hence, we have

$$\Pr[X_c \leq t | x_1 \geq 2R'] = P_2(x_1)(1 - e^{-\frac{\lambda_{ji} t}{2}}) \int_{x_1-2R'}^{\infty} f_c(x)dx, \quad (14)$$



(a) link propagation delay with time-invariant vehicle speed (b) link propagation delay with time-variant vehicle speed (c) link transfer delay without traffic light control (d) link transfer delay with traffic light control

Fig. 6. Validation of the delay analysis.

where $P_2(x_1) = 1 - \int_0^{\frac{R}{v}} \lambda_{ji} e^{-\lambda_{ji}\tau} d\tau \int_{x_1 - R - \tau v}^{\infty} f_c(x)dx$ is the probability that the two clusters cannot be connected directly.

Combining the above two cases and the distribution of x_1, we have

$$\Pr[X_c \le t] = \frac{1 - e^{-\frac{\lambda_{ji} t}{2}}}{e^{-\lambda_{ij}\frac{R}{v}}} \left[\int_{\frac{R}{v}}^{\frac{2R'}{v}} \lambda_{ij} e^{-\lambda_{ij}\tau} P_1(\tau v) d\tau + \int_{\frac{2R'}{v}}^{\infty} \lambda_{ij} e^{-\lambda_{ij}\tau} P_2(\tau v) d\tau \int_{v\tau - 2R'}^{\infty} f_c(x)dx \right]. \tag{15}$$

We can further infer that the PDF of X_c satisfies

$$f_{X_c}(t) = \frac{\frac{\lambda_{ji}}{2} e^{-\frac{\lambda_{ji} t}{2}}}{e^{-\lambda_{ij}\frac{R}{v}}} \left[\int_{\frac{R}{v}}^{\frac{2R'}{v}} \lambda_{ij} e^{-\lambda_{ij}\tau} P_1(\tau v) d\tau + \int_{\frac{2R'}{v}}^{\infty} \lambda_{ij} e^{-\lambda_{ij}\tau} P_2(\tau v) d\tau \int_{v\tau - 2R'}^{\infty} f_c(x)dx \right]. \tag{16}$$

Then, we have

$$\mathbf{E}\{X_c^t\} = \frac{2}{\lambda_{ji} e^{-\lambda_{ij}\frac{R}{v}}} \left[\int_{\frac{R}{v}}^{\frac{2R'}{v}} \lambda_{ij} e^{-\lambda_{ij}\tau} P_1(\tau v) d\tau + \int_{\frac{2R'}{v}}^{\infty} \lambda_{ij} e^{-\lambda_{ij}\tau} P_2(\tau v) d\tau \int_{v\tau - 2R'}^{\infty} f_c(x)dx \right]. \tag{17}$$

Note that when the event X_c happens, the message is forwarded to the opposite direction cluster and then the front neighbor CH in 2δ. Then, the forwarding distance, denoted by d_f, satisfies

$$\begin{aligned} \mathbf{E}\{d_f\} &= v(\mathbf{E}\{X_c^t\} + 2\delta) + \mathbf{E}\{x_1\} + \mathbf{E}_{ij}\{c_s\} \\ &= v(\mathbf{E}\{X_c^t\} + 2\delta) + (R + v/\lambda_{ij}) + \mathbf{E}_{ij}\{c_s\} \end{aligned} \tag{18}$$

where $\mathbf{E}\{X_c^t\} + 2\delta$ is the average time interval for the message being forwarded to the next CH.

The data delivery speed, denoted by v_d, is estimated by

$$\mathbf{E}\{v_d\} = \frac{\mathbf{E}\{d_f\}}{\mathbf{E}\{X_c^t\} + 2\delta} = v + \frac{(R + \lambda_{ij}^{-1}) + \mathbf{E}_{ij}\{c_s\}}{\mathbf{E}\{X_c^t\} + 2\delta}. \tag{19}$$

Finally, we obtain the following theorem to estimate the link propagation delay.

Theorem 5.1. *Let d_{ij} be the distance from intersection i to its neighbor intersection j and X_{ij} be the corresponding propagation delay.*

$$\mathbf{E}\{X_{ij}\} \approx \max\left\{ \frac{d_{ij} - \mathbf{E}_{ij}\{c_s\}}{\mathbf{E}\{v_d\}}, 0 \right\}, \tag{20}$$

where $\mathbf{E}\{v_d\}$ can be obtained from (19).

In (20), there is a $-\mathbf{E}_{ij}\{c_s\}$ term, since at the beginning the message can be fast forwarded to a CH.

Also, we use simulation to verify the analysis of the average link propagation delay, where the parameters are the same as those used in Fig. 4 in the previous section. The analytical results and simulation results with 95 percent confidence intervals are shown in Fig. 6a, demonstrating that our estimation is accurate. Then, considering the dynamic of vehicle speed, we set that the speed of each vehicle will fluctuate in $[0.8v, 1.2v]$ dynamically, and we found that the analytical results and simulation results are still with a high confidence intervals (≥ 90 percent) as shown in Fig. 6b. Using the previous approach in the literature, e.g., (17) in [13], which also estimates the delay considering a similar scenario, when the density is 10 per km, the delay is 36.6 s, much larger than the simulation result (about 5 s).

5.2 Link Transfer Delay

Link Transfer Delay Estimation. We then investigate how long the message can be forwarded to the given direction at the intersection, i.e., the link transfer delay, by taking all possible cases into consideration. The difficulty is to consider the case that the message may travel to the wrong direction and then being carried back until it reaches the right vehicle towards the desired direction at the intersection. Define $\lambda_{es}^n = \lambda_{en} + \lambda_{sn}$. The following theorem provides the estimation of the average link transfer delay and its proof is given in Appendix B, available in the online supplemental material.

Theorem 5.2. *Let Z_{wn}, Z_{en} and Z_{sn} be the link transfer delay from west, east and south to north, respectively, at an intersection. Their average values can be estimated by solving the following equation set:*

$$\begin{bmatrix} \mathbf{E}\{Z_{wn}\} \\ \mathbf{E}\{Z_{en}\} \\ \mathbf{E}\{Z_{sn}\} \end{bmatrix} = \begin{bmatrix} 1 & -B_{wn}^e & -B_{wn}^s \\ -B_{en}^w & 1 & -B_{en}^s \\ -B_{sn}^w & -B_{sn}^e & 1 \end{bmatrix}^{-1} \begin{bmatrix} C_{wn} \\ C_{en} \\ C_{sn} \end{bmatrix}, \tag{21}$$

where w, s, e, n denote directions, and how to obtain the values of the parameters in the right-hand side is given in the proof in Appendix B in the supplementary material, available in the online supplemental material.

Fig. 6c compares the average link transfer delay in simulation with the analytical results (obtained from Theorem

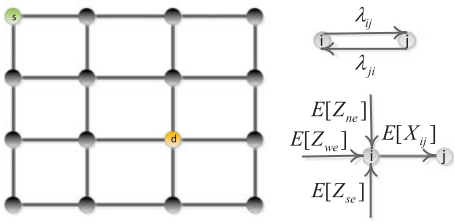


Fig. 7. Weighted graph construction.

5.2). It is observed that the average link transfer delay in theory and in simulation are almost identical. In the worst case, the difference between them is less than 1 s when the vehicle density is 5 per km. The reason for the discrepancy is that when $\lambda = 5$, using $|c_s| = \lceil \mathbf{E}\{c_s\} * \lambda \rceil$ to estimate the number of vehicles in the cluster introduces a relatively large quantization error. In summary, Theorem 5.2 gives an accurate estimation of the average link transfer delay. Then, we consider the traffic light control in an intersection. We set the time of both red light and green light as 10 seconds (for a case study) and the traffic lights change periodically. As shown in Fig. 6d, it is observed that the traffic light will decrease the link transfer delay since the opportunity of that the data carrier can find an appropriate vehicle for the next-hop forwarding in the intersection is enlarged by the traffic light, while the change of delay, w.r.t. the vehicle density has the similar trend and Theorem 5.2 still guarantees a relatively accurate estimation of the delay.

5.3 End-to-End Path Delay

Note that end-to-end path delay is equal to the sum of all link propagation delay and link transfer delay along a path, as concluded in the following theorem.

Theorem 5.3. *Suppose that a message is forwarded according to a path $i_0 \rightarrow i_1 \rightarrow \cdots \rightarrow i_m$, and let $D_{i_0 i_m}$ be the path delay. Then, we have*

$$\mathbf{E}\{D_{i_0 i_m}\} = \sum_{\ell=1}^{m} \mathbf{E}\{X_{i_{\ell-1} i_\ell}\} + \sum_{l=1}^{m-1} \mathbf{E}\Big\{ Z_{\overrightarrow{i_{l-1} i_l i_{l+1}}} \Big\},$$

where $\mathbf{E}\{X_{i_{\ell-1} i_\ell}\}$ and $\mathbf{E}\{Z_{\overrightarrow{i_{l-1} i_l i_{l+1}}}\}$ can be calculated with (20) and (21), respectively.

It should be pointed out that in the above theorem, we use $d_{ij} - R'$ instead of d_{ij} to calculate $\mathbf{E}\{X_{i_{\ell-1} i_\ell}\}$ with (20), since the vehicle carrying the message can start searching for the next-hop carrier before it arrives at the intersection.

5.4 Minimum Delay Routing Algorithm

Given a graph and the vehicle density in each direction of each road segment, we can calculate the average link propagation delay and the link transfer delay for each link and each intersection by Theorems 5.1 and 5.2, respectively. After obtaining these delays, we can set the weights to each link and each intersection from one direction to the others, and then we can construct a weighted graph, and an example is shown in Fig. 7. Based on the weighted graph, we can use the classic Dijkstra's algorithm to find the optimal path with the lowest delay. We design the optimal path selection algorithm called Minimum Delay Routing Algorithm as follows.

The source node uses MDRA to select the path with the minimum expected delay. The selected path is used for routing, and the link forwarding process along the road segments and link transfer process described in the previous section are used as the data propagation strategy by vehicles involved. In a very dynamic network where the vehicle intensity changes drastically and the real-time traffic intensity information can be obtained, it is possible for message carriers to recalculate the optimal path using MDRA, which is a future research issue.

6 Performance Evaluation

To validate the analysis and compare the proposed solution with the existing work, we first simulate the scenario by letting the vehicle inter-arrival process follow the Poisson process as assumed in the analysis, i.e., the vehicles will have an exponential headway distribution. We then use a log-normal headway distribution to investigate the performance of the proposed MDRA routing algorithm, followed by further discussions on the impacts of other assumptions used in our analytical model.

6.1 Exponential Headway Distribution

We compare MDRA with two heavily investigated algorithms, flooding-based Epidemic [25] and geocast-based E-GPSR [26]. Using the flooding-based Epidemic algorithm, whenever two vehicles are within the transmission range of each other, the message will be forwarded to the one who does not have a copy. Therefore, the Epidemic algorithm can guarantee the lowest possible propagation delay with the downside of a very high communication and storage cost, and it is used as a benchmark. Geocast-based routing has been proposed to substantially reduce the communication cost by forwarding the message to the neighbor that is closest to the destination only [28]. Many research works have been done to further improve or optimize the geocast-routing protocols for VANETs [26], [27]. Given the above grid-topology of the VANET, E-GPSR based approaches will select the shortest path in terms of the shortest physical distance.

Algorithm 1. Minimum Delay Routing Algorithm

1: **Input:** Given the source node s and destination node d, graph $G = (V, E)$ and λ_{ij} for each $\langle i, j \rangle \in E$.
2: Run the generalized Dijkstra's shortest-path algorithm, starting from the two-hop neighbors of node d. Calculate the minimum delay by

$$\mathbf{E}\{X^*_{id}\} = \min\Big\{ \mathbf{E}\{X_{ij}\} + \mathbf{E}\{Z_{\overrightarrow{ijj^*}}\} + \mathbf{E}\{X^*_{jd}\} | j \in N_i \Big\}$$

and the corresponding neighbor node

$$i^* = \arg\min\Big\{ \mathbf{E}\{X_{ij}\} + \mathbf{E}\{Z_{\overrightarrow{ijj^*}}\} + \mathbf{E}\{X^*_{jd}\} | j \in N_i \Big\}$$

for each node i ($\in$ V).
3: **Output**: the optimal path from s to d, denoted by p^*_{sd}. For each node i, record $\mathbf{E}\{X^*_{id}\}$ and i^*.

The vehicle communication range is the same as the setting before (200 m). Fig. 8 shows the network graph. The physical distance of each road segment is 100 unit (i.e., $100 \times 20 = 2{,}000$ m). Considering the traffic flow model [29], the vehicle speed v depends on the one-way vehicle arrival

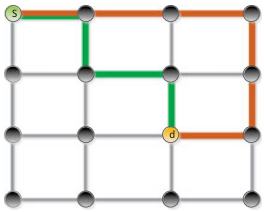
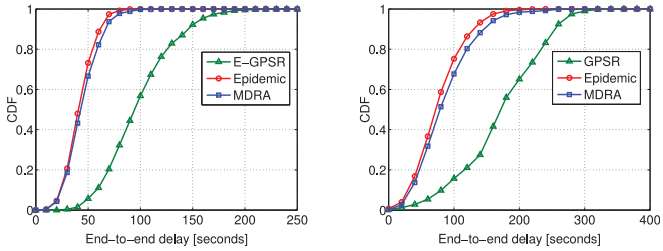


Fig. 8. The optimal path selection: the red path is selected by MDRA and the green path is selected by E-GPSR.



(a) exponential headway (b) log-normal headway

Fig. 9. CDF of the path delay under different headway distributions, simulation results.

rate λ: $v = v_f(1 - \frac{\lambda}{\lambda_{\text{jam}}})$, where v_f is the free-flow speed (equal to the road speed limit), and λ_{jam} is the vehicle arrival rate at which traffic flow comes to a halt. We set $v_f = 20$ m/s and λ_{jam} equal to 100 vehicles per km. Vehicles enter the network from the boundary intersections following Poisson processes, and they will randomly turn to different directions according to pre-set probabilities until they leave the network (e.g., turning east when they enter an intersection in the farthest right column of the graph). We let $\lambda_{ij} = 10$ per km for each direction of each road segment marked in red in Fig. 8, and 5 per km for the rest of the links. This setting mimics the scenario that some road segments (e.g., main street) are more popular and have a higher vehicle density. Flow conservation is guaranteed as the aggregated amount of traffic toward and leaving each intersection equal to each other. We also tune the vehicle turning probabilities at each intersection to ensure the density of each road segment equals the above setting.

Under E-GPSR and MDRA, the green and red path shown in Fig. 8 are selected respectively as the optimal path. In the simulation, we use the same forwarding process as described in Section 5 for both of them.

We conducted 1,000 runs of Monte Carlo simulation for each routing algorithm. In all simulation runs, messages were all successfully delivered to the destination for all routing algorithms considered, which shows that the proposed propagation strategy is effective. The messages experienced different delay using different routing algorithms. Fig. 9a presents the CDF of the path delay. It is observed that the distribution of the delay using MDRA is close to that using Epidemic. For E-GPSR, although the shortest path in terms of the physical distance is selected, it has a much larger delay in most cases. Thus, a path with a longer distance (e.g., the red path selected by MDRA) may have a smaller end-to-end delay for message propagation if the traffic intensity is uneven. Table 2 provides the average path delays using the three algorithms, where the analytical results are calculated using Theorem 5.3. It is found that our theorem can obtain an accurate delay estimation, where the error between simulation and theorem is smaller than 2 percent. Then, we consider the estimation errors of the vehicle density of the road segment. We use $\lambda(1 \pm r * 10\%)$ (where r is a random number selected in [0, 1]) as the real density in the simulation. As shown in the last two columns of Table 2, both the average delay and the delay deviation are affected by the vehicle density estimation error slightly. Thus, the path selected based on the analysis and the proposed algorithm can have a close-to-optimal performance so long as the estimation error is bounded. Further, given the non-uniform vehicle intensity setting, the average delay under MDRA is only about a half of that under E-GPSR.

6.2 Lognormal Headway Distributions

In the previous sections, the theoretical results are obtained with the assumption that the vehicle arrival follows a Poisson process (i.e., the exponential distribution of the inter-vehicle distance), which is a reasonable assumption for roads with relatively low densities according to traffic measurement results [2]. On the other hand, log-normal distribution has been considered a good fit to model the inter-vehicle distances when the vehicle density is high and vehicles need to maintain a safe distance [30], [31], [32]. Here, we use the log-normal distribution to replace the exponential headway distribution in the simulation, and set $\mu = 1$ and $\sigma^2 = 2[-\log(\lambda_{ij}) - 1]$ so the vehicle density remains the same.

Again, for all 1,000 simulation runs for each routing algorithm, messages were successfully delivered to the destination. Fig. 9b compares the delay results using the three routing algorithms. It is observed that MDRA can still have a much smaller delay than E-GPSR, and its corresponding delay distribution is close to that of Epidemic. The average delay of the three algorithms are all increased, since the log-normal distribution has a larger variance than that of exponential distribution which results in a smaller average cluster size. Furthermore, we also use other kinds of distributions for inter-vehicle distances, e.g., Gamma and Gaussian distributions, and the results show similar trends and thus are omitted here. Therefore, even without the assumption that vehicle arrival process is a Poisson process, MDRA still can maintain the performance advantages, e.g.,

TABLE 2
Delay Comparison

Algorithms	Average delay (analysis)	Average delay (simulation)	Delay deviation (simulation)	Average delay with density error (simulation)	Delay deviation with density error (simulation)
MDRA	53.40 s	54.05 s	16.28 s	53.38 s	16.57 s
Epidemic	N/A	52.97 s	15.21 s	51.78 s	14.33 s
E-GPSR	104.32 s	106.50 s	34.11 s	106.67 s	35.65 s

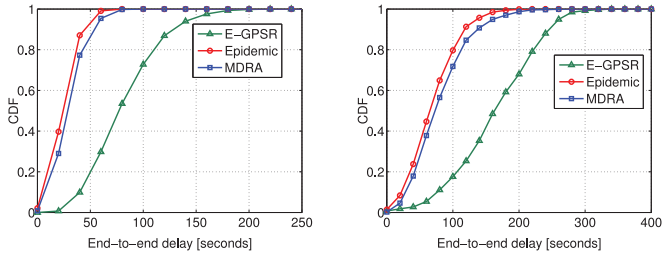


Fig. 10. CDF of the path delay under different headway distributions considering dynamic speed and intersection traffic light, simulation results.

a comparable delay as Epidemic while unicasting a single copy of message only.

Although the proposed algorithm can maintain a good performance, the average delay changes when the headway distributions change. Delay estimation under non-Poisson vehicle arrival processes needs further investigation.

6.3 Impact of Dynamic Vehicle Speed, Vehicle Density, and Traffic Light

We next discuss the impact of the main assumptions used in our model. First, the speed of vehicles may change, so that the clusters are dynamically formulated, and it will also change the formation of clusters. Hence, considering the dynamic vehicle speed, a dynamic cluster management solution should be used [33], [34]. Second, considering the dynamic vehicle density, the average density may not be equal to the real-time density of each road segment, and the performance of the algorithms may change. Third, in city scenarios, when we consider intersections, the traffic light or other traffic control will introduce more uncertainty, and it will affect the delay performance.

From the simulation results given in Section 5, we have found that the dynamic vehicle speed will not change the link delay and link transfer delay substantially, while the traffic light will have a positive impact on the link transfer delay. From Table 2, we have found that the average delay and delay deviation will be slightly affected by the estimation error of vehicle density due to the counteraction of the influences. However, if the average densities of the road segments are always larger (smaller) than their real-time values, then the average delay will be decreased (increased) in the real forwarding process which can be inferred from Fig. 6. Taking both the dynamic vehicle speed and the traffic light into consideration (we use the same parameter settings as in Section 5), we compare the delay results using the three routing algorithms. As shown in Fig. 10, it is observed that the three routing algorithms have a similar performance as that in Fig. 9. However, it should be pointed out, the average delay is decreased for all these three algorithms due to the positive impact of the traffic light.

7 Conclusions

This paper investigated the delay in two-dimensional VANETs using bidirectional vehicles to forward messages based on the proposed propagation strategy. We have developed an analytical framework to accurately estimate the path delay, including the link propagation delay and link transfer delay. Based on the theorems, we further proposed the MDRA algorithm for selecting the path with the lowest expected delay, and the simulation results demonstrate that the proposed MDRA can achieve the delay performance close to the flooding-based Epidemic routing while only a single copy of the message is maintained in the network.

There are many research issues beckoning for further investigation in addition to those listed in the previous sections. Here, we use simple simulation to capture the vehicle traveling behaviour and message propagation delay statistics, and the communication delays are ignored. More sophisticated analysis and experiments using vehicle traffic traces, dynamic network topology, and realistic communication protocols will be valuable to fully understand the end-to-end delay performance including the propagation delay, transmission delay, and packet losses and delay due to wireless network congestion, etc. In this paper, the communication delay is not considered yet since it is much smaller than the carry-and-forward delay by vehicles. For instance, the EU project METIS considers that a maximum delay of 5 ms for one attempt of V2V transmission [35]. However, for some applications, the packet size may be large and the network may suffer from congestion, which lead to a larger communication delay (up to 100 ms), which may not be ignored and can be an interesting aspect for further investigation. Further, the propagation strategy and routing protocol design need to be enriched in handling many realistic situations that are not considered, such as vehicles stop in the middle of a road segment, cluster dynamic due to heterogeneous vehicle speed, etc. Nevertheless, the analytical framework proposed in this paper provides important insights and points out a promising direction on the routing protocol design for large-scale, two-dimensional VANETs.

Acknowledgments

The authors would like to thank the editor and the anonymous reviewers for their helpful and constructive comments, which have helped them improve the manuscript significantly. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Natural Science Foundation of China (NSFC) under grant 61503332 and 61473251.

References

[1] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *Proc. 5th ACM Int. Symp. Mobile ad hoc Netw. Comput.*, 2004, pp. 187–198.

[2] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1538–1556, Oct. 2007.

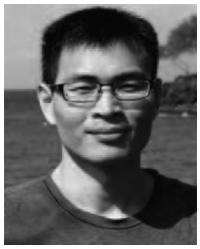
[3] K. Abboud and W. Zhuang, "Modeling and analysis for emergency messaging delay in vehicular ad hoc networks," in *Proc. IEEE Global Telecommun. Conf.*, 2009, pp. 1–6.

[4] J. He, L. Cai, P. Cheng, and J. Pan, "Delay minimization for data dissemination in large-scale VANETs with buses and taxis," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 1939–1950, Aug. 2016.

[5] Y. Bi, H. Shan, X. Shen, N. Wang, and H. Zhao, "A multi-hop broadcast protocol for emergency message dissemination in urban vehicular ad hoc networks," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 3, pp. 736–750, Mar. 2016.

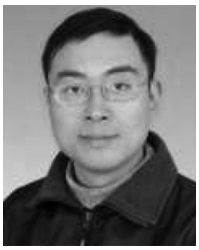
[6] E. Baccelli, P. Jacquet, B. Mans, and G. Rodolakis, "Highway vehicular delay tolerant networks: Information propagation speed properties," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1743–1756, Mar. 2012.

[7] N. Benamar, K. Singh, M. Benamar, D. El Ouadghiri, and J. Bonnin, "Routing protocols in vehicular delay tolerant networks: A comprehensive survey," *Comput. Commun.*, vol. 48, pp. 141–158, 2014.
[8] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3609–3626, Sep. 2009.
[9] A. Abdrabou, B. Liang, and W. Zhuang, "Delay analysis for sparse vehicular sensor networks with reliability considerations," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4402–4413, Sep. 2013.
[10] J. Zhao and G. Cao, "VADD: Vehicle-assisted data delivery in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1910–1922, May 2008.
[11] I. Leontiadis and C. Mascolo, "GeOpps: Geographical opportunistic routing for vehicular networks," in *Proc. IEEE Int. Conf. World Wireless Mobile Multimedia Netw.*, 2007, pp. 1–6.
[12] Y. Liu, J. Niu, J. Ma, L. Shu, T. Hara, and W. Wang, "The insights of message delivery delay in VANETs with a bidirectional traffic model," *J. Netw. Comput. Appl.*, vol. 36, no. 5, pp. 1287–1294, 2013.
[13] H. Saleet, R. Langar, K. Naik, R. Boutaba, A. Nayak, and N. Goel, "Intersection-based geographical routing protocol for VANETs: A proposal and analysis," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4560–4574, Nov. 2011.
[14] Y. Wang, J. Zheng, and N. Mitton, "Delivery delay analysis for roadside unit deployment in intermittently connected VANETs," in *Proc. IEEE Global Telecommun. Conf.*, 2014, pp. 155–161.
[15] M. Khabbaz, C. Assi, and A. Ghrayeb, "Modelling and delay analysis of intermittently connected roadside communication networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2698–2706, Jul. 2012.
[16] J. Jeong, S. Guo, Y. Gu, T. He, and D. Du, "Trajectory-based data forwarding for light-traffic vehicular ad hoc networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 5, pp. 743–757, May 2011.
[17] A. Abdrabou and W. Zhuang, "Probabilistic delay control and road side unit placement for vehicular ad hoc networks with disrupted connectivity," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 129–139, Jan. 2011.
[18] Z. Zhang, G. Mao, and B. D. Anderson, "Stochastic characterization of information propagation process in vehicular ad hoc networks," *IEEE Trans. Intell. Transport. Syst.*, vol. 15, no. 1, pp. 122–135, Feb. 2014.
[19] GB Traffic Volumes, May 2005. [Online]. Available: http://www.mapmechanics.com
[20] F. Bai and B. Krishnamachari, "Spatio-temporal variations of vehicle traffic in VANETs: facts and implications," in *Proc. 6th ACM Int. Workshop Veh. InterNETworking*, 2009, pp. 43–52.
[21] Y. Zhuang, J. Pan, and L. Cai, "A probabilistic model for message propagation in two-dimensional vehicular ad-hoc networks," in *Proc. 7th ACM Int. Workshop Veh. InterNETworking*, 2010, pp. 31–40.
[22] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proc. 10th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2009, pp. 299–308.
[23] S. Shioda, J. Harada, and Y. Watanabe, "Fundamental characteristics of connectivity in vehicular ad hoc networks," in *Proc. IEEE 19th Int. Symp. Pers. Indoor Mobile Radio Commun.*, 2008, pp. 1–6.
[24] K. A. Hafeez, L. Zhao, J. W. Mark, X. Shen, and Z. Niu, "Distributed multichannel and mobility-aware cluster-based MAC protocol for vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3886–3902, Oct. 2013.
[25] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Duke Univ., Durham, NC, Tech. Rep. CS-200006, 2000.
[26] T. Hu, M. Liwang, L. Huang, and Y. Tang, "An enhanced GPSR routing protocol based on the buffer length of nodes for the congestion problem in VANETs" in *Proc. IEEE 10th Int. Conf. Comput. Sci. Educ.*, 2015, pp. 416–419.
[27] F. Granelli, G. Boato, D. Kliazovich, and G. Vernazza, "Enhanced GPSR routing in multi-hop vehicular communications through movement awareness," *IEEE Commun. Lett.*, vol. 11, no. 10, pp. 781–783, Oct. 2007.
[28] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. 6th Annu. Int. Conf. Mobile Comput. Netw.*, 2000, pp. 243–254.
[29] J. D. Fricker and R. K. Whitford, *Fundamentals of Transportation Engineering: A Multimodal Systems Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[30] G. Zhang, Y. Wang, H. Wei, and Y. Chen, "Examining headway distribution models with urban freeway loop event data," *Transport. Res. Record: J. Transport. Res. Board*, no. 1999, pp. 141–149, Jan. 2007.
[31] R. J. Cowan, "Useful headway models," *Transport. Res.*, vol. 9, pp. 371–375, 1975.
[32] G. Yan and S. Olariu, "A probabilistic analysis of link duration in vehicular ad hoc networks," *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 4, pp. 1227–1236, Dec. 2011.
[33] K. Abboud and W. Zhuang, "Stochastic modeling of single-hop cluster stability in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 226–240, Jan. 2016.
[34] C. Wu, X. Chen, Y. Ji, S. Ohzahata, and T. Kato, "Efficient broadcasting in VANETs using dynamic backbone and network coding," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6057–6071, Nov. 2015.
[35] Scenarios, requirements and KPIs for 5G mobile and wireless system, "ICT-317669-METIS/D1.1, METIS deliverable D1.1," Apr. 2013. [Online]. Available: https://www.metis2020.com/documents/deliverables/
[36] M. Ni, J. Pan, L. Cai, J. Yu, H. Wu, and Z. Zhong, "Interference-based capacity analysis for vehicular ad hoc networks," *IEEE Commun. Lett.*, vol. 19, no. 4, pp. 621–624, Apr. 2015.



Jianping He (M'15) received the PhD degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2013. He is currently a research fellow in the Department of Electrical and Computer Engineering, University of Victoria. His research interests include the control and optimization of sensor networks and cyber-physical systems, the scheduling and optimization in VANETs and social networks, and the investment decision in financial market and electricity market. He is a member of the IEEE.

Lin Cai (S'00-M'06-SM'10) received the MASc and PhD degrees in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2002 and 2005, respectively. Since 2005, she has been in the Department of Electrical & Computer Engineering, University of Victoria, where she is currently a professor. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic over wireless, mobile, ad hoc, and sensor networks. She has received the NSERC Discovery Accelerator Supplement Grants in 2010 and 2015, and the best paper awards of IEEE ICC 2008 and IEEE WCNC 2011. She has served as a TPC symposium co-chair for IEEE Globecom'10 and Globecom'13, and the associate editor of the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Vehicular Technology*, the *EURASIP Journal on Wireless Communications and Networking*, the *International Journal of Sensor Networks*, and the *Journal of Communications and Networks*. She is a senior member of the IEEE.



Jianping Pan received the bachelor's and PhD degrees in computer science from Southeast University, Nanjing, China, and he did his postdoctoral research at the University of Waterloo, Waterloo, Canada. He is currently a professor of computer science with the University of Victoria, Victoria, Canada. He also worked with Fujitsu Labs and NTT Labs. His area of specialization is computer networks and distributed systems, and his current research interests include protocols for advanced networking, performance analysis of networked systems, and applied network security. He has been serving on the technical program committees of major computer communications and networking conferences including IEEE INFOCOM, ICC, Globecom, WCNC, and CCNC. He is the Ad Hoc and sensor networking symposium co-chair of the IEEE Globecom 2012 and an associate editor of the *IEEE Transactions on Vehicular Technology*. He is a senior member of the IEEE.

Peng Cheng (M'10) received the BE degree in automation, and the PhD degree in control science and engineering both from Zhejiang University, Hangzhou, P.R. China, in 2004 and 2009, respectively. Currently, he is an associate professor in the Department of Control Science and Engineering, Zhejiang University. He serves as associate editor of *Wireless Networks*, and the *International Journal of Distributed Sensor Networks*. He is also the guest editor of the *IEEE Transactions on Control of Network Systems*. He served as the publicity co-chair of the IEEE MASS 2013 and local arrangement chair for ACM MobiHoc 2015. His research interests include networked sensing and control, cyber-physical systems, and robust control. He is a member of the IEEE.