Optimal Dropbox Deployment Algorithm for Data Dissemination in Vehicular Networks

Jianping He, *Member, IEEE*, Yuanzhi Ni, *Student Member, IEEE*, Lin Cai, *Senior Member, IEEE*,
Jianping Pan, *Senior Member, IEEE*, and Cailian Chen, *Member, IEEE*

Abstract—For vehicular networks, dropboxes are very useful for assisting the data dissemination, as they can greatly increase the contact probabilities between vehicles and reduce the data delivery delay. However, due to the costly deployment of dropboxes, it is impractical to deploy dropboxes in a dense manner. In this paper, we investigate how to deploy the dropboxes optimally by considering the tradeoff between the delivery delay and the cost of dropbox deployment. This is a very challenging issue due to the difficulty of accurate delay estimation and the complexity of solving the optimization problem. To address this issue, we first provide a theoretical framework to estimate the delivery delay accurately. Then, based on the idea of dimension enlargement and dynamic programming, we design a novel optimal dropbox deployment algorithm (ODDA) to obtain the optimal deployment strategy. We prove that ODDA has a fast convergence speed, which is less than κ ($\kappa < n$) iterations for convergence. We also prove that the computational complexity of ODDA is $O(n\kappa m \log m)$, i.e., ODDA has a polynomial computational complexity for a given m, the number of dropboxes for deployment. Performance evaluation by simulation demonstrates the superior performance of the proposed strategies compared with the benchmark methods.

Index Terms—Vehicular networks, data forwarding, delay minimization, dropbox deployment

◆

1 INTRODUCTION

Vehicular communication networks have attracted great attention in recent years thanks to the development in both mobile communication technology and vehicle technology. With the mobility, flexibility and other characteristics of vehicular networks, many applications can be launched to improve service quality and user experience in daily lives, e.g., the road safety, travel comfort, and trip efficiency, and facilitate commercial advertisements, interactive entertainment, urban sensing, etc. [1]. Data dissemination is a crucial and important issue, on which most of these vehicular network applications depend. For example, in the applications of commercial advertisements, the promotional materials from hotels or shopping malls may be attractive to the taxi or bus passengers around the airport. To realize such kind of services cost-effectively, it is important to disseminate the latest information to the targeted area within a short delay by using the vehicle-to-vehicle communication, which is the typical data dissemination problem. Extensive research works have been done for data dissemination problems in vehicular networks, e.g., [2], [3], [4]. However, how to guarantee the highly effective and reliable data dissemination in large-scale vehicular networks is still an open issue and worth further investigation.

1.1 Motivations

To guarantee the effective and reliable data dissemination in large-scale vehicular networks, many existing works have already used the infrastructures or dropboxes or throwboxes to assist the data dissemination in vehicular ad hoc networks (VANETs), and they have shown that using assisted infrastructures can greatly increase the contact probabilities between vehicles and reduce the delivery delay [3], [4], [5], [6], [7], [8], [9]. For example, Banerjee et al. in [5] first examined the performance-cost trade-offs for VANETs by considering assisted infrastructures. It demonstrated that using a small amount of infrastructures can significantly improve the performance of a routing protocol. Wu et al. in [4] pointed out that it is able to improve the data delivery ratio and reduce the delivery delay in VANETs by deploying infrastructures. He et al. in [3] introduced a store-and-forward framework for VANETs with extra storage using "dropboxes", which functions similar to network routers, to address the scalability and high-mobility issues for data dissemination in VANETs. However, deploying infrastructures/dropboxes is costly [5]. Although usually using more dropboxes will bring more benefits on data dissemination, it also incurs more cost, which implies a trade-off between the performance of data dissemination and the cost of deploying dropboxes. As the delivery delay is the major concern in data dissemination [17], [18], it is natural to consider the trade-off

---

- J. He and C. Chen are with the Department of Automation, Shanghai Jiao Tong University, and the Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China. E-mail: {jphe, cailianchen}@sjtu.edu.cn.
- Y. Ni is with the School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China. E-mail: niyuanzhi@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.

Manuscript received 22 Aug. 2016; revised 10 July 2017; accepted 18 July 2017. Date of publication 31 July 2017; date of current version 2 Feb. 2018.
(Corresponding author: Lin Cai.)
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.2017.2733534

between the delivery delay and the deployment cost. Hence, these works motivated us to consider the dropboxes as an assistant of data dissemination and study the optimal deployment problem.

The deployment problem in vehicular networks has been widely studied in [15], [16], [30], [32], [33], [35]. Most of them focused on the highway or straight road segment scenario aiming at minimizing the delivery delay. However, due to the lack of accurate modeling of the benefit of deploying dropboxes, the performance evaluation is limited, particularly in two-dimensional road grid scenarios. Meanwhile, considering the complexity of solving the deployment problem, the optimal solution is also very difficult to obtain.

Given the above observations, the objective of this paper is to build a precise analytical framework and design an optimal deployment strategy with low-computational complexity for data dissemination in large-scale vehicular networks.

1.2 Challenges

There are two main challenges to solve the deployment problem. First, how to obtain the accurate delay estimation and model of the benefit of deploying a dropbox in two-dimensional vehicular networks. Accurate delay estimation has been recognized as a very challenging issue in vehicular networks due to the dynamic of vehicles, especially in two-dimensional vehicular networks [13], [15], [21]. Second, it is difficult to obtain that how many dropboxes and where they should be deployed. If we take all possible strategies into consideration and compare them exhaustively, the computational complexity will be exponentially increasing with the number of intersections, which is impractical. Hence, it needs to design a low-complexity algorithm to solve this problem for large-scale vehicular networks.

1.3 Contributions

In this paper, we formulate a utility-based maximization problem to solve the above challenges. For the delay estimation, by extending the framework in [36], we further quantify the reduction of delay by deploying a dropbox at an intersection. Then, we use the approach of information dimension enlargement and dynamic programming to design a low complexity algorithm to solve the optimization problem. In summary, the contributions of this paper are listed as follows. First, we provide a theoretical framework to estimate the delivery delay and the benefit of deploying dropboxes in terms of delay reduction. Second, we design a novel ODDA to solve the optimal deployment problem. We also prove that ODDA obtains the optimal solution and has a low-computational complexity (i.e., $O(n\kappa m \log m)$). Finally, simulations have been conducted to validate the correctness of the analysis results and the effectiveness of the proposed algorithm.

1.4 Organization

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 forwarding strategy. Section 4 provides some preliminaries. Section 5 introduces ODDA algorithm. Performance evaluations by simulation are presented in Section 6, followed by the concluding remarks and future research issues in Section 7.

2 RELATED WORK

Delay Analysis. Delay analysis for VANETs in sparse highway scenarios has been studied in the literature [10], [11], [14], [19], [20], [22]. In [10], Wisitpongphan et al. proposed a model to quantify the average packet-delivery delay between disconnected vehicles based on their observation of the vehicle inter-arrival time and inter-vehicle spacing. Also, they suggested a store-carry-forward routing mechanism for VANETs. In [11], Huang analyzed the accurate distribution of the information re-healing delay in a sparse bidirectional highway scenario based on [10]. However, the authors in [10], [11] only considered the information delivery delay in an infrastructure-less sparse highway scenario without roadside units (RSUs). In [22], Khabbaz et al. proposed several information release mechanisms to achieve a delay-minimal information delivery in the context of an intermittent roadside network. [19], [20] investigated the delivery delay given the delivery distance and vehicle density, where a one-dimensional road with bidirectional traffic is considered. [31] provided the accurate modeling and analysis framework to estimate the information propagation speed considering the 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 VANETs remains unsolved. Modeling and analysis of carry-and-forward delay in VANETs are complicated due to the intermittent connectivity, dynamic temporal motions, and random forwarding process. As suggested in [14], how to refine the existing models or to design new ones for accurate delay estimation has been investigated in [36].

However, all of them considered how to send the information to passing vehicles, rather than to infrastructures by passing vehicles. Based on [10], Reis et al. in [30] developed a mathematical model to analyze the re-healing time, which is the time required to deliver information between a couple of source and destination nodes in a two-way highway scenario in the presence of infrastructures acting as relays, both connected and disconnected. The results show that the re-healing time is significantly reduced in the presence of infrastructures, in particular for the case of connected infrastructures.

Deployment Problem. In [16], Abdrabou et al. presented a mathematical model for calculating the message delivery delay distribution on a two-lane road. The original packets are carried by the generator until meeting an infrastructure, and the copies of each packet are also sent to carrier vehicles in the opposite direction which may leave the road randomly. However, they considered a scenario where infrastructures are connected through fiber or broadband wireless links. [15] proposed a multi-hop packet delivery delay analytical framework for RSU deployment satisfying a probabilistic delay requirement in a road segment for a low-density VANET. They focused on the vehicle-to-infrastructure packet delivery analysis and obtained the minimum number of RSUs required to cover a straight road segment. [32] analyzed the average information delivery delay for a message to be delivered to both neighbor RSUs in order to obtain the proper deployment distance between two neighbor RSUs under a specific delay constraint. Sun et al. [33] presented an RSU deployment scheme for onboard units (OBUs) to update their short-time certificates

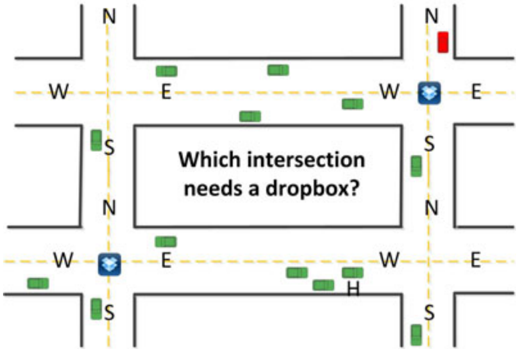


Fig. 1. Scenario.

within the time constraint including the certain driving time (DT) and the extra overhead time. [35] proposed a cluster-based RSU deployment scheme with distributed power control aiming at maximizing the propagation performance and minimizing the power consumption in highway.

In the above works, they took the vehicle speed, the vehicle density, the distance between infrastructures and other randomness in vehicular network message delivery into account for infrastructure deployment. However, they focused on the straight road segment or one-dimensional deployment problem where the two-dimensional delivery path from the source to the destination is not considered. Considering the high cost of the deployment, deploying infrastructures on each road segment may be too expensive. Moreover, with limited deployment resources, how to maximize the utility of the deployment cannot be addressed using existing methods, either. Thus, an analytical framework for deployment from the global, two-dimensional network point of view is needed. [34] developed an Internet gateway deployment technique and a packet routing scheme to guarantee in-vehicle Internet access via multi-hop communications. The objective of the paper is to minimize the deployment cost subject to the required probability of the vehicle connecting to the Internet. The total utility of the deployment is not considered in [34].

In our work, the accurate delay analysis with dropboxes is provided along with the optimal deployment strategy considering both the utility and the cost.

3 SYSTEM MODEL AND FORWARDING STRATEGY

3.1 Scenario and Assumptions

Scenario. In this paper, we consider a data dissemination problem in vehicular networks (e.g., the advertisement of a hotel being delivered to the vehicles near the airport). Suppose there are dropboxes (we use dropboxes for representing of the assisted infrastructure) deployed at the road intersections to help the data dissemination. Given the source region and the destination region, we use the carry-and-forward mechanism for the data dissemination, which is widely used in [3]. When the data is forwarded to an intersection and there is a dropbox at the intersection, the vehicle carrying the message will forward the data to the dropbox first, and then the dropbox will find an appropriate next-hop vehicle for data dissemination based on the designed delivery strategy. It is preferable for data dissemination in vehicular networks when the dropbox has Internet access. However, note that Internet access for dropboxes is highly costly and may not be possible when the Internet connection is damaged [17]. Some application messages are effective only in given areas, e.g., advertisements, and they do not require Internet access. Hence, the dropboxes in our scenario are adopted as the relays, and the Internet is not available for the dropboxes. Then, the objective of this paper is to select the intersections to deploy the dropboxes and the path for data forwarding, to achieve an optimal trade-off between delivery delay and cost. Fig. 1 shows an example of the scenario.[1]

Assumptions. We assume that vehicles can discover others within their communication range (denoted by R) through periodic beacon messages. Assume that 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 [17]. Vehicles are also assumed to be pre-loaded with digital map (e.g., MapMechanics [23]), which provides the off-line street-level map. The message delivery information, e.g., the locations of the source and destination, message expiration time, etc., are assumed to be specified by the source and placed in the packet header. We assume that each dropbox can communicate with a vehicle when the distance between them is not larger than R_d.

3.2 Traffic Model and Communication Model

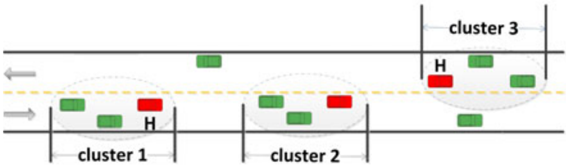
Traffic Model. We abstract a vehicular network to a weighted directed graph $\mathcal{G} = (\mathcal{V}, E, \mathcal{W})$ where $\mathcal{V}$ is the set of nodes denoting the road intersections, and E is the set of links (edges) denoting road segments between two intersections, and $\mathcal{W}$ is the set of weights denoting the average vehicle arrival rate in each link (per second). The value of each element in $\mathcal{W} = \{\lambda_{ij}\}_{n\times n}$ can be estimated from the traces in history [25], [27]. By referring to [17], [25], [26], the vehicles traveling from one intersection to the other can be approximated by a Poisson process, with the average arrival rate $\lambda_{ij} \geq 0$ for $\langle i,j\rangle \in E$. Thus, the inter-vehicle arrival times have an i.i.d. Exponential distribution. We assume that the vehicles have the same speed, denoted by v_{ij}, in the same direction of a road segment $\overrightarrow{ij}$. The vehicle density (the number of vehicles per meter) is λ_{ij}/v_{ij}. This traffic model can be used to describe the characteristics of urban vehicular networks [25], [28].

V2V Communication. Suppose that two vehicles can communicate with each other through a short-range wireless channel. A cluster is defined as the maximal set of vehicles moving in the same direction of a road segment in which every pair of vehicles is connected by at least one single/multi-hop path. The foremost vehicle in a cluster is the cluster head, denoted by H, see Fig. 2a as an example for illustration. The vehicles in the same direction and within each other's communication range will self-organize to become a

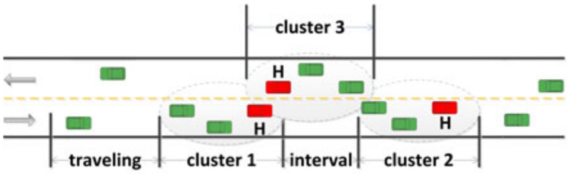
---

[1] There are many available communication systems that can support vehicular communications, e.g., WiFi, cellular communication technologies, etc. Due to the various service requirements from emerging applications in vehicular networks, the existing communication infrastructure may not always have sufficient coverage, capacity, service guarantee, and/or reasonable cost [24]. In this work, we mainly focus on the dropbox deployment algorithm design in hybrid V2V/V2I networks, which do not specify the communication technologies.

(a) Cluster, cluster head, and cluster size



(b) Initially cluster 1 and cluster 2 are disconnected



(c) Cluster 1 and cluster 2 are connected through cluster 3 in the opposite direction after traveling for a time period

Fig. 2. Vehicle cluster and message forward process.

TABLE 1
Important Notations

Symbol	Definition
R	the communication range of vehicles
R'	the effective communication range of two clusters being connected
R_d	the communication range of dropboxes
λ_{ij}	the arrival rate from intersection i to j
λ_{sn}	the arrival rate from south to north at an intersection
T_{sd}	the delivery delay from s to d
c_s	the cluster size
v_{ij}	the vehicle speed in link $\langle i, j\rangle$
X_{ij}	the link propagation delay of $\langle i, j\rangle$
Z_{wn}	the link transfer delay from west to north
κ	the maximum hops among all the paths from s to d
$\mathbf{E}$	the expectation of random variables
δ	the required minimum time of two clusters to exchange the message
$f_c(x)$	the probability distribution function of cluster size
$\overrightarrow{ij}$	the direction from intersection i to j

cluster, while it should be pointed out that the opposite-direction vehicles will not be included in the same cluster. The cluster information is shared by all vehicles in the cluster. Since all vehicles in the same direction of a road segment are assumed to have the same speed, the clusters remain unchanged until they reach an intersection (the work [29] studied how to maintain or reorganize the cluster considering the network dynamics which is beyond the scope of this work). Assume that the message exchange time by wireless communications between vehicles in the same cluster is negligible.

V2I Communication. Suppose that when vehicles come into the communication range of a dropbox at an intersection, they can forward the data to the dropbox or receive the data packet transmitted from the dropbox. The communication between vehicles and dropboxes is named V2I communication.

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

3.3 Forwarding Strategy

There are two parts of the forwarding: how the message is forwarded along a road segment, named the link forwarding strategy, and how the message arriving at an intersection is forwarded to a vehicle traveling in a given direction, named the intersection forwarding strategy.

Link Forwarding Strategy. For the link forwarding strategy, we use the same as that in [36]. Specifically, we assume that the connectivity information is shared within each cluster. The cluster head, H, will broadcast beacon messages (which include the cluster size) periodically to probe for communication opportunities with other Hs in the opposite direction. When these two cluster heads are within the range of R, they can communicate and exchange message within a time interval of δ which depends on the beacon interval, message size, and transmission rate, etc. As shown in Fig. 2b, cluster 1 and cluster 2 are two disconnected clusters traveling in the same direction. Cluster 3 is another vehicle cluster traveling in the opposite direction. When the head of cluster 3 encounters the head of cluster 2, they can learn each other's connectivity information. After traveling for a period of time, as illustrated in Fig. 2c, when the head of cluster 3 encounters the head of cluster 1, they exchange their connectivity information, and find out that cluster 3 can connect with cluster 1 and cluster 2 simultaneously. Using this strategy, data can be forwarded from cluster 1 to cluster 2 with the assistance of cluster 3 during the connected period.

Intersection Forwarding Strategy Without Dropboxes. When an H carrying the message arrives at a tagged intersection (i.e., the intersection is within the communication range of the H), it begins to search for the vehicle traveling to the direction following the selected path and then forwards the message to it. If the data carrier failed to find an appropriate vehicle before it leaves the intersection, it 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 carry the message and forward it to the next vehicle encountered which is 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 selected direction at the tagged intersection, so the message can follow the selected path until reaching the destination.

Intersection Forwarding Strategy with Dropboxes. When H carrying the message arrives at a tagged intersection where a dropbox is deployed, it forwards the message to the dropbox when it is within the communication range of the dropbox. After the dropbox receives the message, it begins to search for the vehicle traveling to the direction following the selected path and then forwards the message to it.

3.4 Problem of Interest

Define Ω to be the directed path set from a source node s to the destination node d, i.e.,

$$\Omega = \{p_{sd}^k | k = 1, 2, \ldots, M\},$$

where p_{sd}^k[2] is one path and M the total number of possible paths. Define Θ to be the dropbox deployment strategy set, which is given by

$$\Theta = \{\mathbf{s} | s_i = 0 \text{ or } 1, s_i \in \mathbf{s}, \mathbf{s} \in R^{1\times n}\}, \tag{1}$$

where $s_i = 1$ or $s_i = 0$ means that the intersection i has a dropbox or not.

We introduce a utility function, $U(t)$, to describe the benefit of reducing the delivery delay. Assume that $U(t)$ is a continuous and differentiable convex function, and is decreasing with t which means that a longer delay yields a lower utility. Let C_d be the cost of deploying a dropbox. Considering the trade-off between delivery delay and the cost, we formulate a utility-based optimization problem as follows,

$$\max_{\mathbf{s}\in\Theta, p_{sd}^k \in \Omega} U(\mathbf{E}\{T_{sd}(k)\}) - \sum_{i=1}^{n} s_i C_d, \tag{2}$$

where $T_{sd}(k)$ is the delivery delay from s to d under path p_{sd}^k, which aims to find an optimal path and an optimal dropbox deployment strategy such that the total utility of delivery delay minus the cost of dropboxes is maximized.

To solve problem (2), there are two key issues: i) how to select the path for data dissemination, i.e., the routing problem, and ii) which intersections should deploy the dropboxes, i.e., the deployment strategy. These two issues are coupled with each other since the deployment strategy will affect the optimal routing and also the optimal deployment depends on the optimal routing, which makes the problem challenging. Specifically, when the deployment strategy is fixed, problem (2) is simplified to a routing problem. In this case, the main challenge is how to estimate the delay $\mathbf{E}\{T_{sd}(k)\}$ accurately, since the traditional shortest path-based approach can be used to solve this routing problem when $\mathbf{E}\{T_{sd}(k)\}$ is given. Then, we consider the possible deployment strategies. Note that each s_i could be 0 and 1, which means that there are 2^n deploy strategies. Hence, it is hard to solve the problem (2) by comparing the value of the total utility under the optimal routing and the different deployment strategies.

Therefore, there are two interesting while challenging problems to solve problem (2). First, how to estimate the delivery delay $\mathbf{E}\{T_{sd}(k)\}$. Second, how to design an algorithm with low-computational complexity (e.g., solve in polynomial time) to solve the problem. Note that if we can solve problem (2) in polynomial time when the number of the dropboxes, m, is fixed, then problem (2) can be solved in polynomial time due to $m \in [0, n]$. Hence, in the remaining of this paper, we will consider the following simplified optimization problem, with a known m

$$\begin{aligned} &\max_{\mathbf{s}\in\Theta, p_{sd}^k \in \Omega} U(\mathbf{E}\{T_{sd}(k)\}) - mC_d \\ &s.t. \quad \sum_{i=1}^{n} s_i = m. \end{aligned} \tag{3}$$

Clearly, in the above problem, the number of dropboxes is fixed. Thus, the number of the possible strategies is reduced to C_n^m, still with a high complexity.

4 PRELIMINARIES

In this section, we provide the preliminaries, including the cluster size estimation, link propagation delay estimation and link transfer delay (without and with dropbox cases) estimation. It should be pointed out that when the dropbox is not considered, there are many works, e.g., [19], [20], [21], [22], [31], can be used to estimate the link propagation delay approximately. In [36], both the link propagation delay and link transfer delay without dropboxes have been estimated accurately and the detailed analysis is provided. Thus, we refer to it and only give the main results of them to make the paper self-contained. Then, we will focus on the detailed analysis of the delivery delay with dropboxes.

4.1 Existing Results

Cluster Size. Let c_s denote the cluster size, which is a random variable. The following lemma is used to estimate the expected cluster size and its probability distribution function (PDF), and its proof is referred to [10], [26].

Lemma 4.1. *Suppose that the vehicle density in a link $\langle i, j\rangle$ is $\tilde{\lambda}_{ij} = \frac{\lambda_{ij}}{v_{ij}}$ and the inter-vehicle distance follows an i.i.d. exponential distribution. Then, we have*

$$\mathbf{E}\{c_s\} = [1 - (\tilde{\lambda}_{ij}R + 1)e^{-\tilde{\lambda}_{ij}R}]/(\tilde{\lambda}_{ij}e^{-\tilde{\lambda}_{ij}R}), \tag{4}$$

and the PDF of c_s can be approximated as

$$f_c(x) = (x^{k-1}e^{-\frac{x}{\theta}})/(\theta^k\Gamma(k)), x > 0, \tag{5}$$

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 $\mathbf{E}\{c_s^2\}$ is given by (5) in [26].

Link Propagation Delay Estimation [36]. *Link propagation delay* measures the time duration for a message passing through a road segment (between two neighbor intersections). It mainly depends on the time for a message to move from one H to the other H or to the end of the link. Define X_c^t the time needed for a message being held by an H until it reaches the head of its front neighbor cluster with the help of the cluster in the opposite direction. The message propagation process (moving from one H 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 propagation delay, so we give their expected values first

$$\begin{aligned} \mathbf{E}\{X_c^t\} = \frac{2}{\lambda_{ji}e^{-\lambda_{ij}\frac{R}{v_{ij}}}} \Bigg[&\int_{\frac{R}{v_{ij}}}^{\frac{2R'}{v_{ij}}} \lambda_{ij}e^{-\lambda_{ij}\tau}P_1(\tau v_{ij})d\tau \\ &+ \int_{\frac{2R'}{v_{ij}}}^{\infty} \lambda_{ij}e^{-\lambda_{ij}\tau}P_2(\tau v_{ij})d\tau \int_{v_{ij}\tau - 2R'}^{\infty} f_c(x)dx \Bigg], \end{aligned}$$

where $P_1(x_1) = e^{-\lambda_{ji}\frac{2R' - x_1}{v_{ji}}}$, R' is the effective communication range of two clusters being connected, and

$$P_2(x_1) = 1 - \int_0^{\frac{R}{v_{ij}}} \lambda_{ji}e^{-\lambda_{ji}\tau}d\tau \int_{x_1 - R - \tau v_{ij}}^{\infty} f_c(x)dx.$$

---

[2] It usually does not have any direct cycles in each path p_{sd}^k.

Then, the delivery speed, denoted by $v^d_{\overrightarrow{ij}}$, is estimated by

$$\mathbf{E}\{v^d_{\overrightarrow{ij}}\} = v_{ij} + \frac{(R + v_{ij}/\lambda_{ij}) + \mathbf{E}_{ij}\{c_s\}}{\mathbf{E}\{X_c^t\} + 2\delta}. \tag{6}$$

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

Theorem 4.2. *Let d_{ij} be the link propagation distance from intersection i to its neighbor intersection j and X_{ij} be the corresponding propagation delay*

$$\mathbf{E}\{X_{ij}\} = \max\left\{\frac{d_{ij} - \mathbf{E}_{ij}\{c_s\}}{\mathbf{E}\{v^d_{\overrightarrow{ij}}\}}, 0\right\}, \tag{7}$$

where $\mathbf{E}\{v^d_{\overrightarrow{ij}}\}$ can be obtained from (6).

Link Transfer Delay Estimation Without Dropboxes [36]. *Link transfer delay* measures the time duration from the moment a message carrier enters the intersection (distance R away from the intersection) to the moment that the message reaches a vehicle moving towards the target direction. Define $\lambda^n_{es} = \lambda_{en} + \lambda_{sn}$, where λ_{en} and λ_{sn} are the arrival rate from east and south direction to the north direction, respectively. The following theorem is given to estimate the average link transfer delay.

Theorem 4.3. *Let Z_{wn}, Z_{en} and Z_{sn} be the link transfer delay from the west, east and south direction to the north direction, respectively, at an intersection. Their average values can be estimated by the following equation set:*

$$\begin{bmatrix} \mathbf{E}\{Z_{wn}\} \\ \mathbf{E}\{Z_{en}\} \\ \mathbf{E}\{Z_{sn}\} \end{bmatrix} = \begin{bmatrix} 1 & -B^e_{wn} & -B^s_{wn} \\ -B^w_{en} & 1 & -B^s_{en} \\ -B^w_{sn} & -B^e_{sn} & 1 \end{bmatrix}^{-1} \begin{bmatrix} C_{wn} \\ C_{en} \\ C_{sn} \end{bmatrix}, \tag{8}$$

where w, s, e, n denote directions, and how to obtain the values of the parameters in the right-hand side is given [36].

To prove the above theorem, we first suppose that $\mathbf{E}\{Z_{en}\}$ and $\mathbf{E}\{Z_{sn}\}$ are given. We then analyze the relationship among $\mathbf{E}\{Z_{wn}\}$, $\mathbf{E}\{Z_{en}\}$ and $\mathbf{E}\{Z_{sn}\}$, and obtain

$$\mathbf{E}\{Z_{wn}\} - B^e_{wn}\mathbf{E}\{Z_{en}\} - B^s_{wn}\mathbf{E}\{Z_{sn}\} = C_{wn}.$$

Then, changing the direction variable, we can obtain all the relationships among $\mathbf{E}\{Z_{wn}\}$, $\mathbf{E}\{Z_{en}\}$ and $\mathbf{E}\{Z_{sn}\}$, which form a linear matrix equation. Solving this matrix equation, we can obtain (8).

4.2 Link Transfer Delay Estimation with Dropboxes

With the help of a dropbox, the vehicle carrying the message can forward the message to the dropbox directly when the vehicle is within the communication range of the dropbox. Then, the dropbox will search for a suitable vehicle for the next-hop message carrier. We investigate how long the dropbox can find the vehicle traveling to the direction following the selected path, i.e., the link transfer delay with a dropbox. There are two cases for the link transfer delay with a dropbox. First, there are one or more vehicles traveling in the desired direction and within the communication range of R_d when the dropbox receives the message, and then the dropbox will forward the message to the vehicle which is in front of them directly. In this case, the link transfer delay is 0 since the message is directly forwarded to the vehicle traveling to the direction following the selected path without any waiting time. The probability of this case occurring equals the probability that there is a vehicle traveling to the given direction and within the communication range of the dropbox. Hence, we have (here we use the north direction as an example)

$$\int_0^{\frac{R_d}{v_n}} \lambda^n_{esw} e^{-\lambda^n_{esw}\tau} d\tau = 1 - e^{-\lambda^n_{esw}\frac{R_d}{v_n}}, \tag{9}$$

where $\lambda^n_{esw} = \lambda_{en} + \lambda_{sn} + \lambda_{wn}$ and v_n is the vehicle speed in the north direction. Second, if there are no vehicles traveling to the given direction and within the communication range of R_d, then the dropbox needs to wait and search for the first vehicle coming to this direction for the next-hop data carrier. In this case, the link transfer delay should be equal to the waiting time. Then, the corresponding PDF of the waiting time should satisfy $\lambda^n_{esw}e^{-\lambda^n_{esw}t}$. Hence, the expectation of the link transfer delay under this case is $\frac{1}{\lambda^n_{esw}}$. Combining two cases together, we have that the expectation of the link transfer delay with a dropbox, denoted by $\mathbf{E}\{\tilde{Z}_{wn}\}$, satisfies

$$\begin{aligned} \mathbf{E}\{\tilde{Z}_{wn}\} &= 0 \times (1 - e^{-\lambda^n_{esw}\frac{R_d}{v_n}}) + e^{-\lambda^n_{esw}\frac{R_d}{v_n}} \times \frac{1}{\lambda^n_{esw}} \\ &= \frac{1}{\lambda^n_{esw}} e^{-\lambda^n_{esw}\frac{R_d}{v_n}}. \end{aligned} \tag{10}$$

Clearly, $\mathbf{E}\{\tilde{Z}_{wn}\}$ is a decreasing function of $\frac{\lambda^n_{esw}}{v_n}$ and R_d, which means that both a larger vehicle density and a larger communication range of the dropbox benefit the link transfer of the message dissemination.

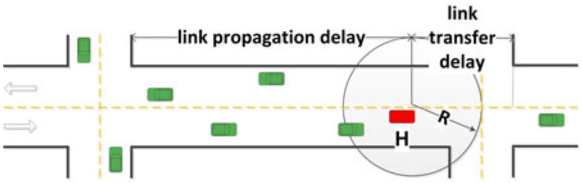
5 OPTIMAL DROPBOX DEPLOYMENT ALGORITHM

In this section, we will design a low-complexity optimal dropbox deployment algorithm (ODDA) to solve problem (3). We first analyze the benefit of the dropbox (i.e., saving the link transfer time), and then present the details of ODDA. Lastly, we provide the detailed performance analysis of ODDA to show the convergence and the complexity of the proposed algorithm.

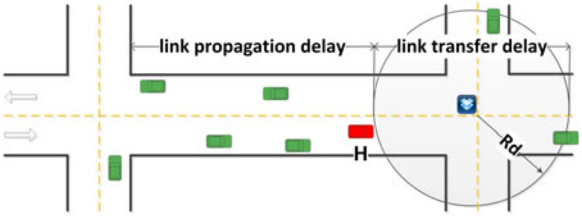
5.1 Delay Reduction of the Dropbox

In this section, we investigate the benefit of using a dropbox at an intersection in terms of delay reduction. Note that with a dropbox, not only the link transfer delay will be decreased, but also the link propagation delay can be decreased. See Fig. 3 as an example for illustration. If there is a dropbox deployed at an intersection, it is observed from Figs. 3a and 3b that the link propagation distance following the direction of the message is decreased due to $R_d \geq R$, and also observed from Figs. 3a and 3c that the link propagation distance of the next-hop of the message would be reduced since the dropbox can forward the message to the vehicle traveling to the given direction and within the communication range of it instantly. Hence, we give the following theorem to estimate the average reduction on link transfer delay of using a dropbox.

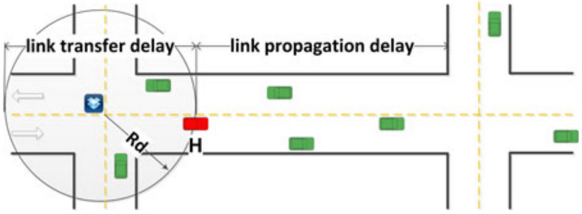
Without loss of generality, we consider the message is meant to be transferred from west to north. When the



(a) Delivery delay without dropboxes



(b) Delivery delay with dropboxes: case 1



(c) Delivery delay with dropboxes: case 2

Fig. 3. Delivery delay with and without dropboxes.

directions are changed, the corresponding results can be easily obtained by changing direction indexes.

Theorem 5.1. *Suppose that a message is needed to be forwarded from direction w to direction n at an intersection. Let B_{wn} be the reduction on link transfer delay of using a dropbox. Then, we have*

$$\begin{aligned}\mathbf{E}\{B_{wn}\} = {} & \frac{(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})}{v^d_n} \\ & \times \left[(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})(R_d-\frac{v_n}{\lambda^n_{esw}})+R_d e^{-\lambda^n_{esw}\frac{R_d}{v_n}}\right] \\ & + \left(\mathbf{E}\{Z_{wn}\}-\mathbf{E}\{\tilde{Z}_{wn}\}\right)+\frac{R_d-R}{v^d_w},\end{aligned} \tag{11}$$

where v^d_w and v^d_n are the delivery speed in the west direction and north direction, respectively.

Proof. There are two cases for the whole link transfer process when we consider the benefit B_{wn}.

First, the dropbox cannot find a suitable vehicle instantly when it receives the message. The probability of this case happens is $e^{-\lambda^n_{esw}\frac{R_d}{v_n}}$. In this case, the message delivery process at the intersection is that the dropbox receives the message from a vehicle in the message arrival direction, which will save a part of the link propagation delay. Then, the dropbox waits and forwards the message to the first vehicle which will travel to the north, so reducing the link transfer delay. Hence, the expected total link transfer time reduction (equals the link propagation delay reduction plus the link transfer delay reduction) in this case satisfies

$$-\frac{R_d-R}{v^d_w}+\frac{1}{\lambda^n_{esw}}, \tag{12}$$

where $-\frac{R_d-R}{v^d_w}$ is the reduced link propagation delay following the message direction.

Second, the dropbox can find a suitable vehicle directly when it receives the message. The probability of this case is $1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}}$. In this case, the link transfer delay is 0, while in both the in and out direction of the message reducing the link propagation delays. Then, the expected total link transfer time reduction in this case satisfies

$$-\frac{R_d-R}{v^d_w}-\frac{\int_0^{\frac{R_d}{v_n}}\lambda^n_{esw}e^{-\lambda^n_{esw}t}(R_d-v_nt)dt}{v^d_n}. \tag{13}$$

By combining the above two cases, the expectation of the B_{wn} can be calculated by

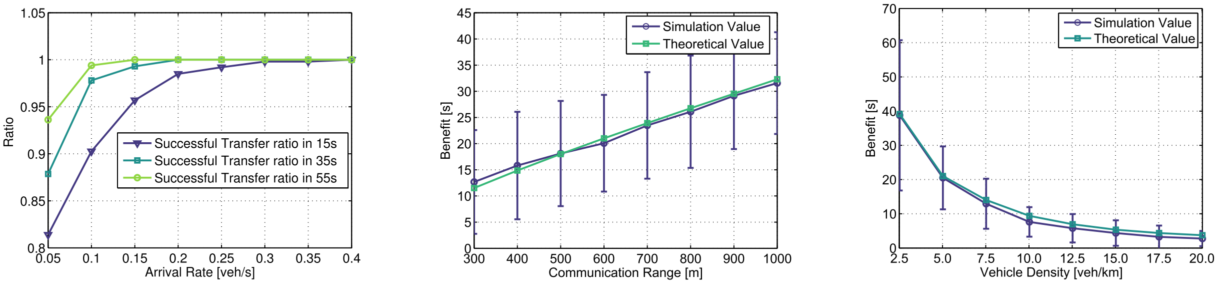
$$\begin{aligned}\mathbf{E}\{B_{wn}\} = {} & \mathbf{E}\{Z_{wn}\}-e^{-\lambda^n_{esw}\frac{R_d}{v_n}}\left[-\frac{R_d-R}{v^d_w}+\frac{1}{\lambda^n_{esw}}\right] \\ & +(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})\frac{\int_0^{\frac{R_d}{v_n}}\lambda^n_{esw}e^{-\lambda^n_{esw}t}(R_d-v_nt)dt}{v^d_n} \\ & +(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})\frac{R_d-R}{v^d_w}.\end{aligned}$$

Note that $\mathbf{E}\{\tilde{Z}_{wn}\}=\frac{1}{\lambda^n_{esw}}e^{-\lambda^n_{esw}\frac{R_d}{v_n}}$, we have

$$\begin{aligned}\mathbf{E}\{B_{wn}\} = {} & \mathbf{E}\{Z_{wn}\}-\mathbf{E}\{\tilde{Z}_{wn}\}+\frac{R_d-R}{v^d_w} \\ & +(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})\frac{\int_0^{\frac{R_d}{v_n}}\lambda^n_{esw}e^{-\lambda^n_{esw}t}(R_d-v_nt)dt}{v^d_n} \\ = {} & \frac{R_d-R}{v^d_w}+\mathbf{E}\{Z_{wn}\}-\mathbf{E}\{\tilde{Z}_{wn}\}+\frac{(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})}{v^d_n} \\ & \times\left[(1-e^{-\lambda^n_{esw}\frac{R_d}{v_n}})(R_d-\frac{v_n}{\lambda^n_{esw}})+R_de^{-\lambda^n_{esw}\frac{R_d}{v_n}}\right],\end{aligned}$$

which completes the proof. □

To validate the above analysis, the simulation of comparing the message delivery performance with or without dropboxes is conducted, where the parameters are given in Section 6. In the simulation, the message is carried by a vehicle traveling towards the intersection and the vehicles on the road segments are randomly generated according to different densities. The message is meant to be transmitted to a targeted direction with or without the help of dropboxes. The successful transfers within three delay bounds are recorded. Besides, the comparisons of different dropbox communication ranges and vehicle densities are provided. Fig. 4 shows the benefit of deploying dropboxes at the intersections. In Fig. 4a, the ratios of successful transfers without dropbox to successful transfers with a dropbox using three different delay bounds are provided. The curves demonstrate that the deployment of dropboxes can effectively help the message transfer when the vehicle arrival rate is rather low or the delay bound is low. Figs. 4b and 4c show the analytical and simulation results of the benefit variation trends with the increment of the communication range and the arrival rate of the vehicles, respectively. As shown in Fig. 4b, the benefit increases with the increment of the



(a) The successful transfer ratio between with and without a dropbox.

(b) The benefit under different communication range settings.

(c) The benefit under different arrival rate settings.

Fig. 4. The advantages of using a dropbox.

communication range of the dropbox because the dropbox can search in a larger area and forward the message to a further away vehicle. Fig. 4c shows when the arrival rate of the vehicles is low, the benefit is significant. The analytical and simulation results in Figs. 4b and 4c are close to each other which validate the analysis.

5.2 Algorithm Design

In this section, we introduce ODDA in details. There are three facts used to design ODDA. The first fact is that if we deploy m dropboxes at the intersections (or nodes thereafter) along with a given path, the optimal strategy is that the m nodes which can get the largest benefit should be chosen for the deployment. The second fact is that if node i is a node in the optimal path p^*_{sd} and there are $\tilde{m}$ dropboxes deployed in the path segment from i and d, then these $\tilde{m}$ dropboxes deployment should be the optimal deployment considering node i as the source node and $\tilde{m}$ as the number of the dropbox deployment. Both of these two facts are easy to be proved by contradiction, and thus is omitted here. The last fact is that the benefit of deploying dropboxes in nodes are independent due to the Poisson vehicle arrival process at each intersection.[3] It means that the delivery delay from each node i to destination node d will not be affected by the previous forwarding and deployment strategies from nodes s to i. Based on these three facts, we employ the idea of dimension enlargement and dynamic programming to design ODDA.

Before giving the details of ODDA, we provide some important notations. Define the information set for each node i by

$$\Phi_i = \begin{bmatrix} Y_i^m & i_m^* & I_i^m \\ Y_i^{m-1} & i_{m-1}^* & I_i^{m-1} \\ \cdots & \cdots & \cdots \\ Y_i^0 & i_0^* & I_i^0 \end{bmatrix}, \tag{14}$$

where for $\ell = m, m-1, \ldots, 0$, Y_i^ℓ denote the minimum delay from i to the destination using ℓ dropboxes, and i_m^* is the optimal next-hop node corresponding to Y_i^ℓ, and I_i^ℓ is the multi-dimensional index vector. I_i^ℓ is given by

$$\begin{aligned} I_i^\ell &= [I_i^\ell(1), \ldots, I_i^\ell(\ell)] = \begin{bmatrix} S_i^\ell \\ \bar{B}_i^\ell \end{bmatrix} \\ &= \begin{bmatrix} i_\ell^1 & i_\ell^2 & \ldots & i_\ell^\ell \\ \mathbf{E}\{B^{i_\ell^1}\} & \mathbf{E}\{B^{i_\ell^2}\} & \ldots & \mathbf{E}\{B^{i_\ell^\ell}\} \end{bmatrix}, \end{aligned} \tag{15}$$

where $S_i^\ell = [i_\ell^1, i_\ell^2, \ldots, i_\ell^\ell]$ is the strategy set, denoting the nodes which have deployed the dropboxes, and $\bar{B}_i^\ell = [\mathbf{E}\{B^{i_\ell^1}\}, \mathbf{E}\{B^{i_\ell^2}\}, \ldots, \mathbf{E}\{B^{i_\ell^\ell}\}]$ is the expected benefit set, denoting the benefit obtained from using dropboxes.

Given $\ell = 0, 1, \ldots, m$, we define three important notations to calculate the delivery delay under different ℓ dropbox deployment strategies by considering the message forwarded from node i to node j and then to node d. Let

$$\mathbf{E}_{ij}^\ell(0) = \mathbf{E}\{X_{ij}\} + \mathbf{E}\{Z_{\overrightarrow{ijj_\ell^*}}\} + \mathbf{E}\{Y_j^\ell\}, \tag{16}$$

where $Z_{\overrightarrow{ijj_\ell^*}}$ is the link transfer delay from direction $\overrightarrow{ij}$ to direction $\overrightarrow{jj_\ell^*}$, be the delivery delay under the strategy that there is no dropbox deployed at node j and node j uses the strategy S_j^ℓ. Let

$$\begin{aligned} \mathbf{E}_{ij}^\ell(1) &= \mathbf{E}\{X_{ij}\} + \mathbf{E}\{Z_{\overrightarrow{ijj_\ell^*}}\} + \mathbf{E}\{Y_j^\ell\} \\ &\quad - \mathbf{E}\{B_{\overrightarrow{ijj_\ell^*}}\} + \min\{\mathbf{E}\{B^{j_\ell^l}\} | B^{j_\ell^l} \in I_j^\ell\}, \end{aligned} \tag{17}$$

where $B_{\overrightarrow{ijj_\ell^*}}$ is the link transfer delay reduction from direction $\overrightarrow{ij}$ to direction $\overrightarrow{jj_\ell^*}$, be the delivery delay under the strategy that a dropbox is deployed at node j and removed one dropbox from the strategy S_j^ℓ. Let

$$\mathbf{E}_{ij}^\ell(2) = \mathbf{E}\{X_{ij}\} + \mathbf{E}\{Z_{\overrightarrow{ijj_{\ell-1}^*}}\} + \mathbf{E}\{Y_j^{\ell-1}\} - \mathbf{E}\{B_{\overrightarrow{ijj_{\ell-1}^*}}\}, \tag{18}$$

be the delivery delay under the strategy that a dropbox is deployed at node j and node j uses the strategy $S_j^{\ell-1}$. For the above three notations, $\mathbf{E}\{X_{ij}\}$, $\mathbf{E}\{Z_{\overrightarrow{ijj_\ell^*}}\}$ and $\mathbf{E}\{B_{\overrightarrow{ijj_\ell^*}}\}$ can be calculated by (7), (8) and (11), respectively.

Remark 1. When the next hop of node j is switched, the values of (16), (17), and (18) will change. Then, we can obtain the minimum delay and the corresponding next hop of node j by comparing their values. For simplicity, we suppose that j_ℓ^* (or $j_{\ell-1}^*$) is the updated optimal next hop according to the strategy used in (16), (17), and (18) in the remainder of this paper.

---

[3] Although the benefits are independent, the minimum delay from s to d depends on both the path selection and deployment strategy which are coupled with each other. Thus, simply choosing the nodes with the largest delay reduction to deploy the dropboxes may not be able to obtain the optimal solution. We have included this approach in the performance comparison in Section 6.

Then, we describe ODDA in Algorithm 1, and its detailed analysis will be given in the following section.

Algorithm 1: Optimal Dropbox Deployment Algorithm

```
1: Input: Given the nodes $s$ and $d$ and graph $G = (\mathcal{V}, E, \mathcal{W})$.
   Let the initial information set satisfy
       $\Phi_d(0) = \begin{bmatrix} \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing \\ \cdots & \cdots & \cdots \\ 0 & \varnothing & \varnothing \end{bmatrix}$    (19)
   and set $\Phi_i(0) = \varnothing$ (all elements are $\varnothing$) for $i \in \mathcal{V}$ and $i \neq d$.
2: Loop
3:   At each iteration, each node sends its information set to
     its neighbor nodes.
4:   If node $i$ receives the information set from its neighbors,
     for $\ell = m, m-1, \ldots, 0$, it calculates
         $\mathbf{E}\{Y_i^\ell\} = \min\{\mathbf{E}_{ij}^\ell(0), \mathbf{E}_{ij}^\ell(1), \mathbf{E}_{ij}^\ell(2) | j \in N_i\}$,    (20)
     where $N_i$ is the neighbor set of node $i$, and then finds the
     optimal next-hop node $i_\ell^*$ by
         $i_\ell^* = \arg\min\{\mathbf{E}_{ij}^\ell(0), \mathbf{E}_{ij}^\ell(1), \mathbf{E}_{ij}^\ell(2) | j \in N_i\}$.
5:   If $\mathbf{E}_{ii_\ell^*}^\ell(0) = \mathbf{E}\{Y_i^\ell\}$,
         $I_i^\ell = I_{i_\ell^*}^\ell$;
     and if $\mathbf{E}_{ii_\ell^*}^\ell(1) = \mathbf{E}\{Y_i^\ell\}$, let
         $I_i^\ell = I_{i_\ell^*}^\ell; I_i^\ell(l) = \left[ \mathbf{E}\{B \xrightarrow[ii_\ell^*[i_\ell^*]_\ell^*]{i_\ell^*} \} \right]$;
     otherwise,
         $I_i^\ell(l) = \left[ I_{i_\ell^*}^{\ell-1}(l), \left[ \mathbf{E}\{B \xrightarrow[ii_\ell^*[i_\ell^*]_{\ell-1}^*]{i_\ell^*} \} \right] \right]$
6:   Node $i$ updates its information set $\Phi_i$.
7:   If the iteration time $t \leq n-1$, go to step 2.
8: End Loop
9: Output: $Y_s^m$ and $I_s^m$. Then, the nodes in the strategy set $S_s^m$
   are where the $m$ dropboxes should be deployed.
```

5.3 Convergence and Complexity of ODDA

In this section, we will prove the convergence of ODDA, and also analyze the computational complexity of the proposed algorithm.

Let $\kappa = \max_{p_{sd}^k \in \Omega} |p_{sd}^k|$ be the maximum number of hops among all the paths from nodes s to d. Then, we give the following theorem, which guarantees that ODDA can converge in less than κ $(\leq n-1)$ iterations and obtain the optimal solution of problem (3).

Theorem 5.2. *Given m dropboxes for deployment, ODDA obtains the optimal deployment strategy $S_s^m(\kappa)$ and the corresponding delay is $Y_s^m(\kappa)$, where $\kappa \leq n-1$ is the iteration index, i.e., ODDA solves the problem (3).*

Proof. Suppose that p_{sd}^k (denoting $s = i_0 \to i_1 \to \cdots \to i_\nu \to d$) is one of the paths in Ω. Let N_i^l be the l-hop neighbor set of node i.

At iteration $t = 1$, each 1-hop neighbor i of node d will update its information set to $\Phi_i(1)$ based on the received information set $\Phi_d(0)$. From the step 4 in Algorithm 1, we have $Y_i^0(1) = \mathbf{E}\{X_{id}\}$. Clearly, we have that

$$\mathbf{E}\{X_{i_\nu d}\} \geq Y_{i_\nu}^0(1),$$

which means that the expected delay from i_ν to d is not smaller than $Y_{i_\nu}^0(1)$.

At iteration $t = 2$, each 2-hop neighbor node, named node j, of node d will update its information set and obtain $Y_j^1(2)$. From the step 4 in Algorithm 1, $Y_j^1(2)$ satisfies

$$Y_j^1(2) = \min\left\{\mathbf{E}_{ji}^1(0), \mathbf{E}_{ji}^1(1), \mathbf{E}_{ji}^1(2) | j \in N_i^1\right\}.$$

Hence, we infer that

$$\mathbf{E}\{X_{i_{\nu-1}d}^1\} \geq Y_{i_{\nu-1}}^1(2),$$

where $X_{i_{\nu-1}d}^1$ denotes the delivery delay under the path $i_{\nu-1} \to i_\nu \to d$ and one dropbox is deployed along this path. We thus conclude that the expected delivery delay under the path $i_{\nu-1} \to i_\nu \to d$ and when one dropbox is deployed along the path is not smaller than $Y_{i_{\nu-1}}^1(2)$.

By a similar analysis, we can obtain that

$$\mathbf{E}\{X_{i_{\nu-m}d}^m\} \geq Y_{i_{\nu-m}}^m(m+1),$$

where $X_{i_{\nu-m}d}^1$ denotes the delivery delay using the path $i_{\nu-m} \to i_\nu \to d$ with m dropbox along this path. It means that the expected delivery delay using the path $i_{\nu-m} \to \cdots \to i_\nu \to d$ with m dropboxes along this path is no smaller than $Y_{i_{\nu-m}}^m(m+1)$ obtained by node $i_{\nu-m}$ at iteration $m+1$.

At iteration $t = m+2$, according to the information set updating, node $i_{\nu-m-1}$ will obtain $Y_{i_{\nu-m-1}}^m(m+2)$. From step 4 in Algorithm 1, $Y_{i_{\nu-m-1}}^m(m+2)$ satisfies

$$\begin{aligned} Y_{i_{\nu-m-1}}^m(m+2) = \min\{&\mathbf{E}_{i_{\nu-m-1}i_{\nu-m}}^m(0), \mathbf{E}_{i_{\nu-m-1}i_{\nu-m}}^m(1), \\ &\mathbf{E}_{i_{\nu-m-1}i_{\nu-m}}^m(2) | j \in N_i^m\}. \end{aligned}$$

Based on the definitions of $\mathbf{E}_{ij}^\nu(0)$, $\mathbf{E}_{ij}^\nu(1)$, and $\mathbf{E}_{ij}^\nu(2)$, one infers that

$$\mathbf{E}\{X_{i_{\nu-m-1}d}^m\} \geq Y_{i_{\nu-m-1}}^m(m+2),$$

where $X_{i_{\nu-1}d}^1$ denotes the delivery delay using the path $i_{\nu-1} \to i_\nu \to d$ and optimally deploying m dropbox along this path. Hence, the expected delay using the path $i_{\nu-m} \to \cdots \to i_\nu \to d$ and optimally deploying m dropboxes along the path is no smaller than $Y_{i_{\nu-m-1}}^m(m+2)$.

Similarly, one can infer that

$$\mathbf{E}\{X_{i_0 d}^m\} \geq Y_{i_0}^m(\nu+1),$$

i.e., the expected delay under using the path $i_0 \to \cdots \to i_\nu \to d$ and deploying m dropboxes along the path is not smaller than $Y_{i_0}^m(\nu+1)$ which is obtained by node i_0 at iteration $\nu+1$.

Since p^k_{sd} could be any path in Ω, we have

$$\mathbf{E}\{X^m_{i^k_0 d}\} \geq Y^m_{i_0}(|p^k_{sd}|), \tag{21}$$

holds for $k = 1, 2, \ldots, M$, where $|p^k_{sd}|$ is the length of p^k_{sd}. Let the minimum delivery delay under the optimal deployment of m dropboxes, be $\mathbf{E}\{X^{m*}_{sd}\}$. Then, $\mathbf{E}\{X^{m*}_{sd}\}$ should satisfy

$$\mathbf{E}\{X^{m*}_{sd}\} = \min\Big\{\mathbf{E}\{X^m_{i^k_0 d}\} | k = 1, 2, \ldots, M\Big\}.$$

From the definition of Y^m_s,

$$Y^m_s(\kappa) = \min\Big\{Y^m_{i_0}(|p^k_{sd}|) \,|\, k = 1, 2, \ldots, M\Big\},$$

is the delivery delay corresponding to the optimal deployment of m dropboxes along a given path.

Therefore, on the one hand, we have $Y^m_s(\kappa) \geq \mathbf{E}\{X^{m*}_{sd}\}$ since $\mathbf{E}\{X^{m*}_{sd}\}$ is the minimum delay under the optimal deployment. On the other hand, from (21), one infers that

$$\begin{aligned} &\min\Big\{\mathbf{E}\{X^m_{i^k_0 d}\} | k = 1, 2, \ldots, M\Big\} \\ &\geq \min\Big\{Y^m_{i_0}(|p^k_{sd}|) \,|\, k = 1, 2, \ldots, M\Big\}, \end{aligned} \tag{22}$$

which means that $\mathbf{E}\{X^{m*}_{sd}\} \geq Y^m_s(\kappa)$. Hence, we have

$$Y^m_s(\kappa) \geq \mathbf{E}\{X^{m*}_{sd}\} \geq Y^m_s(\kappa),$$

i.e.,

$$Y^m_s(\kappa) = \mathbf{E}\{X^{m*}_{sd}\}.$$

Thus, we have completed the proof. □ □

Theorem 5.3. *Suppose that the number of neighbors of each node is no more than a constant, which means that the number of the paths in each intersection is no more than a constant. Then, the computational complexity of ODDA is $O(n\kappa m \log m)$ for the convergence of the algorithm.*

Proof. From Theorem 5.2, we know that after $|p_{sd}|_{\max}$ iterations, ODDA will converge. Meanwhile, note that at each iteration, each node needs to calculate the value of $\mathbf{E}^\ell_{ij}(0)$, $\mathbf{E}^\ell_{ij}(1)$ and $\mathbf{E}^\ell_{ij}(2)$ for $j \in N_i$, and the minimum value of them. For each $\mathbf{E}^\ell_{ij}(0)$ and each $\mathbf{E}^\ell_{ij}(2)$, the computational complexity is 1. For each $\mathbf{E}^\ell_{ij}(1)$, the computational complexity is $O(\log \ell)$ since it needs to calculate the minimum value $\min\{\mathbf{E}\{B^{j^l_\ell}\}|B^{j^l_\ell} \in \bar{B}^\ell_j\}$. Since the number of neighbors of each node is no more than a constant, the computational complexity is $O(1)$ to obtain the minimum value among $\mathbf{E}^\ell_{ij}(0)$, $\mathbf{E}^\ell_{ij}(1)$ and $\mathbf{E}^\ell_{ij}(2)$ for $j \in N_i$. Hence, combining all cases of them, the computational complexity for each node at each iteration is

$$\sum_{\ell=0}^{m} 2 + O(\log \ell) + O(1) = O(m) + O(\log m!) = O(\log m!).$$

Note that the total number of the network nodes is n and it needs κ iterations for ODDA to converge, the total computational complexity satisfies

$$n\kappa O(\log m!) = O(n\kappa m \log m),$$

which has completed the proof. □

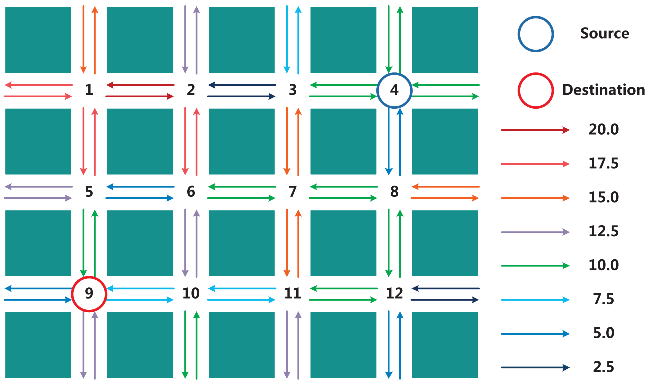


Fig. 5. An example of the two-dimensional road scenario.

6 PERFORMANCE EVALUATION

To validate the analysis and compare the proposed solution with the benchmark solutions, extensive simulation is conducted in this section.

6.1 Parameter Setting

In the simulation, all intersections are four-way, and the road topology is shown in Fig. 5. The length of each road segment is 2 km. The vehicles traveling on the road with a given speed 20 m/s. The vehicle densities of different roads are randomly selected (from 2.5 vehicles/km to 20 vehicles/km) while reflecting the reality [9], [10], [12], [19]. We assume that the message source is at the northeast corner of these intersections and the destination is at the southwest corner, e.g., #4 and #9 in Fig. 5, respectively. From the real-world traffic trace analysis, it follows that the inter-vehicle spacing is exponentially distributed [10]. Thus, the vehicle inter-arrival process follows the Poisson process as assumed in the analysis. The simulation is conducted by Matlab. The message size and data rate are set as constants. For tractability, the delay of each message handover is set to 50 ms, which is typically sufficient for vehicles to handle channel access, data packet transmission, and MAC-layer retransmission operations.

6.2 Deployment Method Comparison

Different deployment methods are compared in this section using the example shown in Fig. 5. The densities of the roads are marked with different colors. The message is generated at #4 intersection and is meant to be delivered to #9 intersection. M is the number of the deployed dropboxes.

As shown in Table 2, we compare the performance of the exhaustive search method, ODDA, shortest path deployment strategy (SP), greedy-based strategy (Greedy) and the highest benefit deployment strategy (H-benefit). For exhaustive search method, we try all possible paths and deployment strategies to obtain the optimal solution. For the shortest path deployment, we first find the path with minimum expected delay ignoring dropboxes. Then the dropboxes with highest benefits are deployed at the intersections of this path. For the greedy-based scheme, the path from the source to the destination is chosen based on a greedy selection principle, i.e., when the message carrier

TABLE 2
Comparison of Deployment Methods

Method	M = 0					M = 1				
	Path	Dep.	Delay (Exp)	Delay (Log)	Delay (Flow)	Path	Dep.	Delay (Exp)	Delay (Log)	Delay (Flow)
Optimal	4-3-7-6-2-1-5-9	n/a	52.5	48.9	53.5	4-3-7-6-10-9	10	40.1	36.7	40.5
ODDA	4-3-7-6-2-1-5-9	n/a	52.5	48.9	53.5	4-3-7-6-10-9	10	40.1	36.7	40.5
SP	4-3-7-6-2-1-5-9	n/a	52.5	48.9	53.5	4-3-7-6-2-1-5-9	7	42.6	40.4	45.2
Greedy	4-3-7-11-10-9	n/a	58	55.2	61.0	4-3-7-11-10-9	10	43.1	43.2	48.0
H-benefit	n/a	n/a	n/a	n/a	n/a	4-3-2-1-5-9	3	80.6	69.3	82.5
Method	M = 2					M = 3				
	Path	Dep.	Delay (Exp)	Delay (Log)	Delay (Flow)	Path	Dep.	Delay (Exp)	Delay (Log)	Delay (Flow)
Optimal	4-3-7-11-10-9	10,11	28.2	30.1	35.7	4-3-7-11-10-9	3,10,11	22.3	26.2	30.6
ODDA	4-3-7-11-10-9	10,11	28.2	30.1	35.7	4-3-7-11-10-9	3,10,11	22.3	26.2	30.6
SP	4-3-7-6-2-1-5-9	5,7	32.7	32.2	36.7	4-3-7-6-2-1-5-9	3,5,7	26.8	28.4	31.7
Greedy	4-3-7-11-10-9	10,11	28.2	30.1	35.7	4-3-7-11-10-9	10,11,3	22.3	26.2	30.6
H-benefit	4-3-2-6-5-9	3,6	95.2	80.4	97.0	4-3-2-1-5-6-10-9	3,5,10	99.8	87.5	103.6

arrives at the intersection, it first searches all the directions towards the destination, and then chooses the direction with a larger vehicle density. For the highest benefit deployment method, we first exhaustively search all the paths and deploy dropboxes in order to achieve the highest benefit on each path. Then the path with the minimum expected delay among those paths with the highest benefit is chosen as the highest benefit deployment solution.

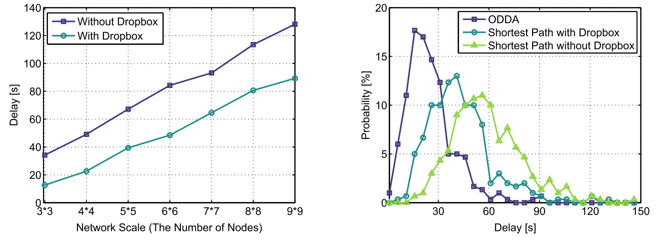
As shown in Table 2, ODDA and the exhaustive search method can achieve the optimal deployment strategy while the shortest path, greedy-based and highest benefit scheme may not be optimal due to the lack of global information and comparison. For the shortest path and greedy-based scheme, the path is selected before deploying dropboxes while the highest benefit scheme only focuses on the benefit of deployed dropboxes. Thus, all these three sub-optimal schemes cannot achieve the globally optimized solution although they may obtain the same output as the optimal solution occasionally. It is worthy to note that the optimal deployment for different number of dropboxes may lead to different delivery paths. In addition to the optimal choice, ODDA can also obtain the solution in a short computation time which is significantly reduced compared with the exhaustive search method.

In addition to the exponential distribution-based inter-vehicle distance as assumed in the analytical framework, we also examine the deployment performance by generating the inter-vehicle distance with log-normal distributions and adjusting the vehicle speed based on the traffic flow model in [40], and conduct 5,000 times Monte Carlo simulations for each setting. This is because the log-normal model and traffic flow model better reflect vehicle interactions in urban scenarios [37], [38], [39], [40]. The mean and the standard deviation of the log-normal distribution is the same as the corresponding exponential distribution. The free-flow speed v_f and traffic jam density $\tilde{\lambda}_m$ used in the traffic flow model (i.e., $v = v_f(1 - \tilde{\lambda}/\tilde{\lambda}_m)$) are 20 m/s and 135 vehicles/km, respectively. As shown in the Delay (Log) and Delay (Flow) columns in Table 2, based on the dropbox deployment strategy of the listed methods, ODDA still achieves the smallest average delivery delay compared with other methods.

Since the vehicle density estimation may be inaccurate and the vehicle velocity may change in reality, comparative simulations are conducted to examine the impacts of the varying vehicle density and the varying velocity. In each road segment, the actual vehicle density is a uniformly distributed random variable ranging from $\pm 10\%$ of its mean value, and the results are given under "Exp ($\triangle\lambda$)" and "Log ($\triangle\lambda$)" in Table 3. The vehicle velocity varies according to a uniformly distributed random variable ranging from $\pm 10\%$ of its mean value in each time slot ("Exp ($\triangle v$)" and "Log ($\triangle v$)"). As

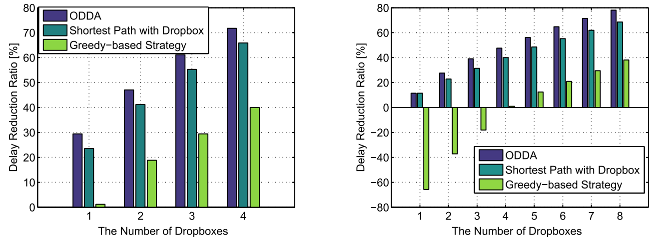
TABLE 3
Comparison of Relaxed-Assumption Cases

Method	M = 0					M = 1				
	Exp ($\triangle\lambda$)	Exp ($\triangle v$)	Log ($\triangle\lambda$)	Log ($\triangle v$)	T-L	Exp ($\triangle\lambda$)	Exp ($\triangle v$)	Log ($\triangle\lambda$)	Log ($\triangle v$)	T-L
ODDA	45.1	48.9	49.3	48.8	46.5	34.3	38.1	36.9	37.3	37.0
SP	45.1	48.9	49.3	48.8	46.5	36.7	41.1	40.5	41.1	38.5
Greedy	52.3	55.7	55.4	55.8	54.0	40.9	43.8	42.9	43.0	40.0
H-benefit	n/a	n/a	n/a	n/a	n/a	70.5	63.7	69.4	64.2	77.5
Method	M = 2					M = 3				
	Exp ($\triangle\lambda$)	Exp ($\triangle v$)	Log ($\triangle\lambda$)	Log ($\triangle v$)	T-L	Exp ($\triangle\lambda$)	Exp ($\triangle v$)	Log ($\triangle\lambda$)	Log ($\triangle v$)	T-L
ODDA	28.4	30.2	30.0	30.8	26.0	24.7	26.4	26.0	26.4	21.0
SP	29.4	33.3	31.8	32.8	28.5	25.4	28.2	28.3	28.6	23.5
Greedy	28.4	30.2	30.0	30.8	26.0	24.7	26.4	26.0	26.4	21.0
H-benefit	84.3	74.4	81.5	75.3	93.0	91.5	82.4	87.6	81.9	96.5



(a) Network scale (b) Delay distribution

Fig. 6. Delay performance comparison.



(a) 3*4 road network (b) 5*6 road network

Fig. 7. Delay reduction ratio comparison.

shown in Table 3, the delivery delays of different relaxed-assumption cases are slightly different but the proposed deployment strategy still guarantees the best delivery performance in terms of the delay given different inter-distance distributions, the varying vehicle density and velocity.

Furthermore, in [36], the impact of independent traffic light control on VANET delay performance has been studied, where at each intersection, traffic light control switches between the green light and the red light periodically. From [36], on average, deploying traffic lights at intersections can slightly reduce the end-to-end delay. This is because, although the message carrier may be delayed at an intersection by the traffic light, it meanwhile may have a better chance to encounter a suitable next-hop carrier or a cluster to forward the message. The negative and positive impacts offset each other and result in a small change in end-to-end delay. The simulation results considering the impact of traffic lights are given under “T-L” in Table 3.

Remark 2. The critical density studied in the percolation theory tells when there exists an infinite-sized cluster in the network, which does not mean that all the vehicles in the network are connected. Whether the vehicle density is above or below the critical density does not change the delay performance derivation. In our simulation, the vehicle density is set in the range from 2.5 to 20.0 vehicles per km, covering the cases that the density is above or below the critical density (12.4 vehicles per km) for percolation, and the simulation results validate that our analysis is accurate in this wide range.

6.3 Delay Evaluation

Fig. 6a shows the delay of messages delivered from the northeast intersection to the southwest intersection in different road network scales. The number of intersections varies from 3*3 to 9*9 and the vehicle density of each road segment is randomly chosen as specified in Section 6.1. For each network scale, we randomly generate 100 different vehicle density distributions and calculate the average message delivery delay from the source to the destination. With the increment of the network scale, the delivery delay without dropboxes increases while deploying dropboxes (using 3 dropboxes for each network scale) can significantly reduce the delivery delay.

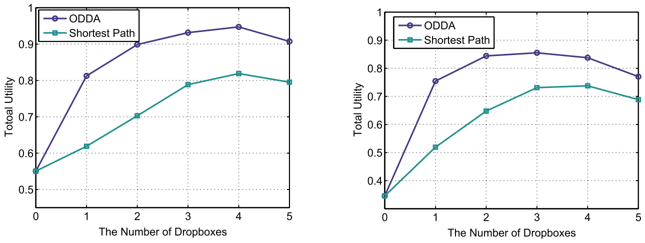
The distributions of the delivery delay using different deployment methods, i.e., ODDA and shortest path with/without dropboxes are compared. Based on a fixed road topology, in each simulation, the vehicles are randomly generated based on the densities of different road segments while the message is delivered along the path selected by each method. We repeat the simulation for 300 times and collect the delivery delay from the source to the destination using the above deployment methods. The collected statistics are divided into 30 intervals as shown in Fig. 6b. It shows that ODDA can achieve a smaller delivery delay compared with the other methods while their distributions have similar trends.

6.4 Improvement with Dropboxes

In this part, the delivery delays with three different deployment methods are compared with the delay along the shortest path without any dropbox, i.e., ODDA, shortest path with dropbox and greedy-based scheme. Fig. 7 shows the delay reduction using these methods in the 3*4 and 5*6 road network, respectively. With the increment of the number of deployed dropboxes, the delivery delay is significantly reduced. The greedy-based method takes the shortest computation time to obtain its solution but the performance is not as good as the other two methods. As shown in Fig. 7b, the greedy-based deployment delay is even larger than the shortest path without any dropbox when the number of the deployed dropboxes is relatively small. This is because sometimes the greedy-based deployment scheme may come up with a path where the delivery delay is much larger than the shortest delivery delay path. ODDA achieves a higher reduction in delay than the shortest path with dropbox strategy while the time complexities are similar. Fig. 7 demonstrates that the proposed ODDA can effectively reduce the delivery delay compared with other sub-optimal deployment methods.

6.5 Utility Evaluation

We choose Gaussian function $U(T) = ae^{(-\frac{(T-b)^2}{2c^2})}$ and quadratic function $U(T) = -dT^2 + e$ as the utility function in (2), respectively, where a, b, c, d and e are constants (we set $a = 2$, $b = -0.1$, $c = 40$, $d = 0.02^2$ and $e = 1.6$). The parameter C_d is set as a constant (we set $C_d = 0.14$). The simulation results of the total utility with different numbers of dropboxes are shown in Figs. 8a and 8b. When the first dropbox is deployed, the total utility raises rapidly. With the increment of the number of the deployed dropboxes, the total utility gradually reaches an upper bound and finally decreases slowly. This is because with the reduction of the delivery delay, the increment of utility function is becoming less obvious/significant while the deployment cost is always increasing linearly. Thus, the total utility first increases and then goes down. From Fig. 8, it is easy to obtain the optimal number of the deployed dropboxes corresponding to the total utility. In addition to that, the total utility of ODDA is always better than other sub-optimal deployment strategies when deploying the same number of dropboxes since the utility function is a decreasing function.



(a) Gaussian function (b) Quadratic function

Fig. 8. Optimization evaluation with different utility functions.

7 CONCLUSION

This paper investigated the optimal dropbox deployment in large-scale vehicular networks considering the trade-off between the benefit and the cost of dropbox deployment. We formulate a utility-based optimization problem to model the problem. To solve this problem, we first provided a complete framework to estimate the delivery delay and calculate the benefit of deploying a dropbox at an intersection. Then, we proposed ODDA to solve the problem. We proved the convergence (optimality) of ODDA and provided the complexity analysis of ODDA. The simulation demonstrated the superior performance of the proposed algorithm compared with the benchmark methods.

Note that the messages would be transmitted between more than one pair of source and destination. In this paper, we focus on the case of a single pair of source and destination and propose an optimal deployment algorithm with a polynomial computational complexity. Obviously, the proposed single-pair solution can be applied when the dropboxes can be incrementally added to support new pairs of sources and destinations, and it can also be extended as a heuristic method to find suitable locations to deploy dropboxes for multi-pair scenarios. However, in this way, the optimality for the multi-pair problem cannot be guaranteed. The global optimal solution for multi-pair scenarios should consider the correlation between each dropbox and the optimal paths corresponding to different pairs of the sources and destinations. How to obtain the optimal deployment strategy for multi-pair scenario requesting further investigation. Furthermore, the impact of vehicle mobility and density on the handover of the message using different MAC protocols and message delivery strategies is an important problem for future research.

ACKNOWLEDGMENTS

The authors would like to thank the editor and the anonymous reviewers for their helpful and constructive comments, which have helped us improve the manuscript significantly. This work was supported in part by the Natural Science Foundation of China (NSFC) under grant U1405251, 61622307 and the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] F. Da Cunha, A. Boukerche, L. Villas, A. Viana, and A. Loureiro, "Data communication in VANETs: A survey, challenges and applications," INRIA Saclay, Palaiseau, France, *Res. Rep.* RR-8498, 2014.

[2] 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.

[3] 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.

[4] Y. Wu, Y. Zhu, and B. Li, "Infrastructure-assisted routing in vehicular networks," in *Proc. IEEE INFOCOM*, 2012, pp. 1485–1493.

[5] N. Banerjee, M. D. Corner, D. Towsley, and B. N. Levine, "Relays, base stations, and meshes: Enhancing mobile networks with infrastructure," in *Proc. 14th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2008, pp. 81–91.

[6] Y. Ding, C. Wang, and L. Xiao, "A static-node assisted adaptive routing protocol in vehicular networks," in *Proc. 4th ACM Int. Workshop Veh. Ad Hoc Netw.*, 2007, pp. 59–68.

[7] Y. Xiong, J. Ma, W. Wang, and J. Niu, "Optimal roadside gateway deployment for VANETs," *Przeglad Elecktrotechnicznyt*, vol. 88, pp. 273–276, 2012.

[8] K. Wong, B. Lee, B. Seet, G. Liu, and L. Zhu, "BUSNet: Model and usage of regular traffic patterns in mobile ad hoc networks for inter-vehicular communications," in *Proc. Int. Conf. Inf. Commun. Technol.*, 2003, pp. 102–108.

[9] M. Xing, J. He, and L. Cai, "Maximum-utility scheduling for multimedia transmission in Drive-Thru Internet," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2649–2658, Apr. 2016.

[10] 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.

[11] J. Huang, "Accurate probability distribution of rehealing delay in sparse VANETs," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1193–1196, Jul. 2015.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] I. Leontiadis and C. Mascolo, "GeOpps: Geographical opportunistic routing for vehicular networks," in *Proc. IEEE Int. Symp. World Wireless Mobile Multimedia Netw.*, 2007, pp. 1–6.

[19] 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.

[20] 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.

[21] Y. Wang, J. Zheng, and N. Mitton, "Delivery delay analysis for roadside unit deployment in intermittently connected VANETs," in *Proc. IEEE Global Commun. Conf.*, 2014, pp. 155–161.

[22] 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.

[23] GB Traffic Volumes, May 2005. [Online]. Available: http://www.mapmechanics.com

[24] National Operations Center of Excellence, "Vehicle to infrastructure deployment coalition," [Online]. Available: https://www.transportationops.org/V2I/V2I-overview, 2016.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] A. B. Reis, S. Sargento, F. Neves, and O. K. Tonguz, "Deploying roadside units in sparse vehicular networks: What really works and what does not," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2794–2806, Jul. 2014.

[31] Z. Zhang, G. Mao, and B. D. Anderson, "Stochastic characterization of information propagation process in vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 122–135, Feb. 2014.

[32] Y. Wang, J. Zheng, and N. Mitton, "Delivery delay analysis for roadside unit deployment in vehicular ad hoc networks with intermittent connectivity," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8591–8602, Oct. 2016.

[33] Y. Sun, X. Lin, R. Lu, X. Shen, and J. Su, "Roadside units deployment for efficient short-time certificate updating in VANETs," in *Proc. IEEE Int. Conf. Commun.*, 2010, pp. 1–5.

[34] H. A. Omar, W. Zhuang, and L. Li, "Gateway placement and packet routing for multihop in-vehicle Internet access," *IEEE Trans. Emerging Topics Comput.*, vol. 3, no. 3, pp. 335–351, Sep. 2015.

[35] J. Tao, L. Zhu, X. Wang, J. He, and Y. Liu, "RSU deployment scheme with power control for highway message propagation in VANETs," in *Proc. IEEE Global Commun. Conf.*, 2014, pp. 169–174.

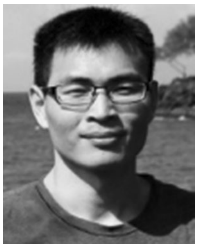
[36] J. He, L. Cai, J. Pan, and P. Cheng, "Delay analysis and routing for two-dimensional VANETs using carry-and-forward mechanism," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1830–1841, Jul. 2016.

[37] G. Zhang, Y. Wang, H. Wei, and Y. Chen, "Examining headway distribution models with urban freeway loop event data," *Transp. Res. Rec.: J. Transp. Res. Board*, vol. 1999, pp. 141–149, 2015.

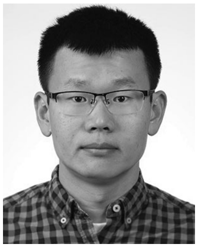
[38] I. Greenberg, "The log-normal distribution of headways," *Australian Road Res.*, vol. 2, no. 7, pp. 14–18, 1966.

[39] G. Yan and S. Olariu, "A probabilistic analysis of link duration in vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1227–1236, Dec. 2011.

[40] J. D. Fricker and R. K. Whitford, *Fundamentals of Transportation Engineering: A Multimodal Systems Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.

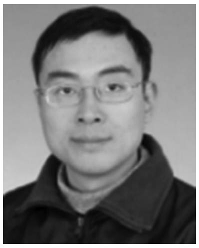


Jianping He (M'15) received the PhD degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2013. He is currently an associate research fellow in the Department of Automation, Shanghai Jiao Tong University. He was a research fellow in the Department of Electrical and Computer Engineering, University of Victoria from 2014 to 2017. His research interests include the control and optimization in cyber-physical systems (CPS), the scheduling and routing in VANETs, and the investment decision in financial market and electricity market. He served as an associate editor for a special issue on consensus-based applications in networked systems in the *International Journal of Robust and Nonlinear Control* in 2016. He is a member of the IEEE.



Yuanzhi Ni (S'17) received the BSc degree in automation from the Nanjing University of Science and Technology, Nanjing, China, in 2012, where he is currently working toward the PhD degree in control theory and engineering. From September 2015 to September 2017, he was a jointly supervised PhD student in the Department of Electrical & Computer Engineering, University of Victoria. His current research interest is vehicular networks. He is a student 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 been a recipient of 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 Globecom10 and Globecom13, and as an 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 (S'96-M'98-SM'08) received the bachelor's and PhD degrees in computer science from Southeast University, Nanjing, Jiangsu, China. He is currently a professor of computer science with the University of Victoria, Victoria, British Columbia, Canada. He did his postdoctoral research with the University of Waterloo, Waterloo, Ontario, Canada. He also worked at 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 received the IEICE Best Paper Award in 2009, the Telecommunications Advancement Foundation's Telesys Award in 2010, the WCSP 2011 Best Paper Award, the IEEE Globecom 2011 Best Paper Award, the JSPS Invitation Fellowship in 2012, the IEEE ICC 2013 Best Paper Award, and the NSERC DAS Award in 2016, and 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 IEEE Globecom 2012 and an associate editor of the *IEEE Transactions on Vehicular Technology*. He is a senior member of the ACM and the IEEE.

Cailian Chen (S'04-M'06) received the BEng and MEng degrees in automatic control from Yanshan University, P. R. China, in 2000 and 2002, respectively, and the PhD degree in control and systems from the City University of Hong Kong, Hong Kong SAR, in 2006. She joined the Department of Automation, Shanghai Jiao Tong University, in 2008 as an associate professor. She is now a full professor. Her research interests include vehicular networks and application in intelligent transportation systems, and industrial wireless sensor and actuator network. She has authored and/or coauthored two research monographs and more than 100 referred international journal and conference papers. She is the inventor of more than 20 patents. She received the "IEEE Transactions on Fuzzy Systems Outstanding Paper Award" in 2008. She won the First Prize of Natural Science Award twice from The Ministry of Education of China in 2006 and 2016, respectively. She was honored Changjiang Young Scholar by Ministry of Education of China in 2015 and Excellent Young Researcher by NSF of China in 2016. She serves as associate editor of the *IEEE Transactions on Vehicular Technology*, and *Peer-to-Peer Networking and Applications* (Springer), and also served as Symposium TPC Co-chair of IEEE Globecom 2016 and VTC2016-fall, and TPC member of several flagship conferences including IEEE Globecom, IEEE ICC, and IEEE WCCI. She is a member of the IEEE.