

**© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.**

# Throughput of Infrastructure-based Cooperative Vehicular Networks

Jieqiong Chen, *Student Member, IEEE*, Guoqiang Mao, *Senior Member, IEEE*,  
Changle Li, *Senior Member, IEEE*, Ammar Zafar, *Member, IEEE*, Albert Y. Zomaya, *Fellow, IEEE*

**Abstract**—In this paper, we provide detailed analysis of the achievable throughput of infrastructure-based vehicular network with a finite traffic density under a cooperative communication strategy, which explores combined use of vehicle-to-infrastructure (V2I) communications, vehicle-to-vehicle (V2V) communications, mobility of vehicles and cooperations among vehicles and infrastructure to facilitate the data transmission. A closed form expression of the achievable throughput is obtained, which reveals the relationship between the achievable throughput and its major performance-impacting parameters such as distance between adjacent infrastructure points, radio ranges of infrastructure and vehicles, transmission rates of V2I and V2V communications and vehicular density. Numerical and simulation results show that the proposed cooperative communication strategy significantly increases the throughput of vehicular networks, compared with its non-cooperative counterpart, even when the traffic density is low. Our results shed insight on the optimum deployment of vehicular network infrastructure and optimum design of cooperative communication strategies in vehicular networks to maximize the throughput.

**Index Terms**—Data dissemination, cooperative communication, throughput, vehicular networks, vehicle-to-infrastructure communications, vehicle-to-vehicle communications.

## I. INTRODUCTION

Vehicular networks have recently gained significant interest from academia and industry because of their increasingly important role in improving road traffic efficiency, enhancing road safety and providing real-time information to drivers and passengers [1]–[3]. By disseminating real-time information about traffic accidents, traffic congestion or obstacles in the road, road safety and traffic efficiency can be greatly improved. Furthermore, offering value-added services like digital maps with real-time traffic status and in-car entertainment services can greatly enhance the convenience and comfort of drivers and passengers.

Vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communications, on one hand, are two fundamental techniques to disseminate data for vehicular applications; on the other hand, as pointed out in our previous paper [4] and other work [5], [6], V2V communications may become unreliable when the number of hops in the communication becomes large, and incur long communication delay when the vehicular density is low. Moreover, V2I communications may have limited availability due to the limited number of infrastructure points attributable to the high deployment cost, especially in rural areas and in the initial deployment phase of vehicular network. Therefore, V2I and V2V communications may have to co-exist and complement each other to meet the

diverse communication requirements of vehicular networks ranging from safety information dissemination to in-car entertainment services.

In this paper, we consider a scenario where there is a vehicle of interest (VoI) wanting to download a large-size file, e.g., a video, from the Internet, and all the other vehicles (termed *helpers*) assist its download using a cooperative communication strategy, which explores the combined use of V2I communication, V2V communication, the mobility of vehicles and cooperation among vehicles and infrastructure to facilitate data transmission. The scenario being considered corresponds to the category of delay-tolerant applications. We are interested in the long-term data rate the VoI can achieve, i.e., the achievable throughput, which is one of the most important performance metrics in wireless (vehicular) networks because it characterizes the feasible data dissemination rate of the network. In our previous work [4], under the same network setting, we have analyzed the achievable throughput of vehicular network with the assumption that the data rate of V2I communications is larger than the data rate of V2V communications. In this paper, we extend to consider a more general scenario without the restriction of the aforementioned assumption and give an accurate analysis on the achievable throughput by the VoI when there is finite vehicular density in the network, and investigate the topological impact on the achievable throughput. Our analytical results shed insight on the optimum deployment of vehicular network infrastructure and the design of optimum cooperative communication strategies in finite vehicular networks to maximize the throughput.

The novelty and major contributions of this paper are summarized as follows:

- 1) a cooperative communication strategy is proposed, which utilizes V2I communications, V2V communications, the mobility of vehicles, and cooperations among vehicles and infrastructure to improve the achievable throughput by the VoI;
- 2) an analytical framework is developed for studying the data dissemination process under our cooperative communication strategy. A closed-form expression of the achievable throughput by the VoI in a vehicular network with a moderate number of vehicles or a finite vehicular density is obtained, which reveals the relationship between the achievable throughput and different parameters such as distance between two neighboring infrastructure points, radio ranges of infrastructure and vehicles, transmission rates of V2I communications and V2V communications and density of vehicles;

- 3) both simulation and numerical results show that the proposed cooperative communication strategy significantly increases the throughput of vehicular networks, compared with its non-cooperative counterpart, even when the traffic density is low.

The rest of this paper is organized as follows: Section II reviews related work. Section III introduces the system model, the proposed cooperative communication strategy and the problem formation. Theoretical analysis of the data dissemination process and the achievable throughput are provided in Section IV. In Section V, we validate the analytical results using simulations and discuss the impact of major performance-impacting parameters. Section VI concludes this paper.

## II. RELATED WORK

Since the seminal work of Gupta and Kumar [7], extensive research has been done to investigate the throughput and capacity of wireless networks. Particularly, Gupta and Kumar showed that the maximum throughput of static wireless networks is  $\Theta(\frac{1}{\sqrt{n}})$  with  $n$  being the number of nodes in the network. In [8], Grossglauser and Tse showed that by leveraging on the nodes' mobility, a per-node throughput of  $\Theta(1)$  can be achieved at the expense of unbounded delay. In [9], Mao et al. presented a simple relationship to estimate the capacity of both static and mobile networks and developed a generic methodology for analyzing the network capacity that is applicable to a variety of different scenarios. Focusing on the capacity of vehicular networks, Wang et al. [10] analyzed the asymptotic uplink throughput scaling law of urban vehicular networks with uniformly distributed RSUs. All of the aforementioned work focused on studying the scaling law of the throughput (capacity) when the number of vehicles or vehicular density is sufficiently large. In this paper, we focus on an accurate analysis, instead of the scaling law, of the throughput of vehicular networks with a moderate number of vehicles or a finite vehicular density where the asymptotic analysis may not apply.

Other work has also investigated the performance of vehicular networks, measured by the downloaded data volume [11], transmission delay [12], communication link quality [13] etc. Among the major techniques to enhance these performance measures, cooperative communications, including both cooperations among vehicles and cooperations among infrastructure points, stands out as a popular and important technique. The following work has investigated cooperative communications among vehicles in vehicular networks. In [11], Zhou et al. introduced a cooperative communication strategy using a cluster of vehicles on the highway to cooperatively download the same file from the infrastructure to enhance the probability of successful download. In [12], Zhu et al. studied using multiple nearby vehicles to collaboratively download data from an RSU and analyzed the average download time using the network coding techniques. In [14], Yan et al. developed a theoretical model to analyze the achievable channel data rate of VANETs for cooperative mobile content dissemination, also assisted by network coding techniques. They focused on the transmission throughput, i.e., the channel data rate in MAC

layer, which is different from this work as we focus on the network achievable throughput. Li et al. [15] proposed a push-based popular content distribution scheme for vehicular networks, where large files are broadcast proactively from a few infrastructure points to vehicles inside an interested area and further disseminated cooperatively among vehicles using V2V communications. In [16], Wang et al. introduced a coalitional graph game to model the cooperations among vehicles and proposed a coalition formation algorithm to implement the cooperation between vehicles for popular content distribution. In addition to cooperations among vehicles, cooperations among infrastructure points can also be achieved by caching different files or different parts of a file in different infrastructure points to help moving vehicles to download from the Internet. In [17], to fully utilize the wireless bandwidth provided by APs, Zhang and Yeo proposed a cooperative content distribution system for vehicles by using cooperative APs to distribute contents to moving vehicles. More specifically, by prefetching different data into the selected APs, vehicles can obtain the complete data from those selected APs when traveling through their coverage areas. In [1] and [18], by utilizing infrastructure cooperation for data dissemination, the authors proposed a cooperative content dissemination scheme in vehicular networks to maximize the downloaded data size [1] and the success probability of download [18] respectively. However, the aforementioned work only considered either cooperations among vehicles or cooperations among infrastructure points. In contrast, our work considers both types of cooperations to maximize the throughput.

There are very limited studies considering both vehicular cooperation and infrastructure cooperation. In [19], Merhad et al. explored cooperations among inter-connected infrastructure and V2V communications to efficiently deliver packets from a source vehicle to vehicles far away. In their work, they focused on designing the optimum routing path to reduce the total delay for delivering a packet from a source to its destination. Our focus is on investigating the achievable throughput in vehicular networks when a VoI requests data from the Internet.

## III. SYSTEM MODEL AND PROBLEM FORMATION

### A. Network Model

We consider a highway with bi-directional traffic flows. The highway is modeled by an infinite line with roadside infrastructure, e.g., RSUs, Wi-Fi APs or LTE base stations, regularly deployed with equal distance  $d$ . The width of a lane is typically small compared with the transmission range of vehicles. Therefore, we ignore the road width and model multiple lanes in the same direction as one lane [20]–[22]. We further assume that all infrastructure points are connected to the Internet through wired or wireless infrastructure.

We adopt a widely used traffic model in highway [6], [22], [23], that in each direction (eastbound and westbound), the distribution of vehicles follows a homogeneous Poisson process with densities  $\rho_1$  and  $\rho_2$  respectively. It follows that the inter-vehicle distances in each direction are exponentially distributed. This exponential inter-vehicle spacing distribution has been supported by some empirical study that it can

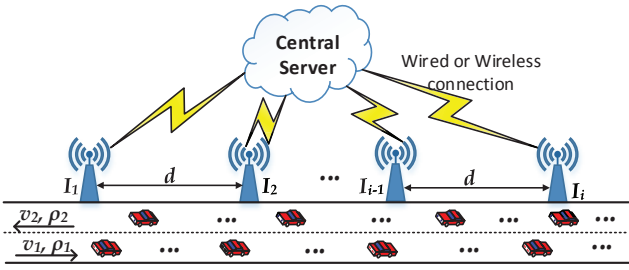


Fig. 1. An illustration of the system model for a bi-directional highway with infrastructure regularly deployed with equal distance  $d$ . The density and speed of vehicles in each direction are  $\rho_1, v_1$  and  $\rho_2, v_2$  respectively.

accurately characterize the real traffic distribution when the traffic density is low or medium [22]. Besides, vehicles in each direction travel at the same constant speed of  $v_1$  and  $v_2$  respectively [6], [22], [24]. In real networks, individual vehicular speed may deviate from the mean speed, e.g., Gaussian speed model [25], [26]. However, such deviations, which results in vehicle overtaking, have only minor impact on the throughput being studied as shown later in our simulation. The system model is illustrated in Fig. 1.

### B. Wireless Communication Model

Both V2I and V2V communications are considered. All infrastructure points have the same radio range, denoted by  $r_I$ ; and all vehicles have the same radio range, denoted by  $r_0$  where  $r_I > r_0$ , which reflects the fact that infrastructure typically has stronger communication capability. A pair of vehicles (or vehicle and infrastructure) can directly communicate with each other if and only if their Euclidean distance is not larger than the radio range  $r_0$  (or  $r_I$ ). There are other more realistic and intricate connection models, e.g., the SINR connection model [27] and the log-normal connection model [28], [29]. This simple unit disk model has been extensively used in the field [6], [21], [30], [31]. It grossly captures the fact that all wireless devices have a limited transmission range and that the closer two devices become, the easier it is for them to establish a connection. This simplification allows us to omit physical layer details and focus on the topological impact of vehicular networks on the throughput, which is the main performance determining factor. We will show later in the simulation that the unit disk model assumption has little impact on the throughput.

We consider that each vehicle has a single antenna so that they can not transmit and receive at the same time. Besides, we adopt unicast transmission that each infrastructure (or vehicle) can only transmit information to one vehicle at a time. Note that in many applications, e.g., sharing emergency or road traffic information among multiple vehicles, broadcast is better used. The scenario being considered in this paper corresponds to a unicast scenario where each user may receive distinct content from the Internet. Both unicast and broadcast are important in vehicular networks [32]. Furthermore, it has been shown in [7] that whether the infrastructure transmit to one vehicle at a time, or divides its bandwidth among multiple users and transmits to multiple users at the same time, does

not affect throughput calculation. Therefore, the assumption that the infrastructure transmits to one vehicle at a time is immaterial to throughput calculation. For the interference model, we assume that V2I and V2V communications are allocated different channels so that there is no mutual interference between them. Besides, we adopt the widely used Protocol Interference Model [33] that a transmitter cannot transmit if there are other transmitter transmitting within its interference range. In our work, the inter-infrastructure distance is large so that infrastructure can transmit simultaneously without causing any mutual interference. For the V2V communications, noting that the VoI is the only receiver of V2V communications, therefore, under the Protocol Interference Model, there will at most one transmitter transmitting its data to the VoI at a randomly chosen time instant, which implies that there will be no interference caused by other simultaneous transmitters. Our work assuming the Protocol Interference Model can be readily extended to another widely used Physical Interference Model (also known as the SINR Model) because it has been established in [33] that any spatio-temporal scheduling satisfying the Protocol Interference Model can also meet the requirement of the corresponding Physical Interference Model when some parameters, like the interference range, transmit power and the SINR threshold are appropriate selected. Furthermore, the MAC protocol associated with the Protocol Interference Model is the CSMA scheme. As we consider a single VoI only in this paper, collisions, which are major concerns of MAC protocol, have little impact on the achievable throughput of the VoI.

We assume V2I and V2V communicate at a constant data rate  $w_I$  and  $w_V$  respectively [7]–[9]. For time-varying channels, the values of  $w_I$  and  $w_V$  can be replaced by the respective time-averaged data rate of V2I and V2V communications and our analysis still applies. Furthermore, differently from our previous paper [4] which assume that  $w_I > w_V$ , in this paper, we remove this assumption and give detailed analysis covering a wider range of scenarios. Indeed, analysis later in this paper will show that depending on the relationship between  $w_I, w_V$  and the speeds of vehicles in both directions, the system can be classified into three regimes: one regime where the throughput is mainly limited by the data rate of V2I communications, i.e.,  $w_I$ ; one regime where the throughput is mainly restrained by the data rate of V2V communications,  $w_V$ ; and another regime where the throughput is determined both by the data rate of V2I communications,  $w_I$ , and the data rate of V2V communications,  $w_V$ . Furthermore, only one-hop communications are considered. This can be explained by the fact that in the specific scenario being considered, there is only one vehicle with download request (VoI), all other vehicles (*helpers*) assist the VoI to receive more data. Any new data in the vehicular network must come from the infrastructure. Therefore, allowing multi-hop V2V communications between the VoI and *helpers*, e.g., allowing V2V communications between *helpers*, only helps to balance the distribution of information among *helpers* but do not increase the net amount of information available in the network. Furthermore, even though allowing more than one hop communication between the VoI and infrastructure is beneficial to the VoI's data downloading because it allows the VoI having longer connection

time with the infrastructure, the improvement is expected to be marginal, especially when the traffic density is small, which has been verified by our simulation result as shown later.

### C. Cooperative Communication Strategy

Now we introduce the cooperative communication strategy considered in this paper. Specially, we consider a scenario where a VoI wants to download a large file, e.g., a video, from a remote server, and analyze the throughput that can be achieved by the VoI via a combined use of V2I communications, V2V communications, vehicular mobility and cooperations among vehicles and infrastructure.

The scenario being considered corresponds to a vehicular network where only a small number of vehicles have requests for large-file downloads. Another scenario that has been widely considered in the literature, commonly known as *the saturated traffic scenario*, considers that all vehicles have requests for download. Saturated traffic scenario is often used in analyzing the capacity of the network [8], [9]. We point out that for the particular problem being considered, i.e., downloading a large file from a remote server, saturated traffic scenario constitutes a trivial case and offer the following intuitive explanation for that. Note that when downloading files from a remote server, the new information (e.g., parts of the files) must come from the infrastructure points. V2V communications only help to balance the distribution of information among vehicles and do not increase the net amount of information available in the system. Therefore, when all vehicles have download requests, it can be easily established that the optimum strategy that maximizes the capacity is for each vehicle to download its requested file directly from the infrastructure. When only single vehicle or a small number of vehicles have download requests, the situation becomes more intriguing. In this situation, other vehicles may help to retrieve information from the infrastructure when these vehicles enter into the coverage of their respective infrastructure points and then deliver the information to the VoI(s) outside the coverage of the infrastructure. In this way, the net amount of new information available in the system is boosted and therefore increasing the throughput (capacity) of the VoI(s).

As mentioned in the beginning of this subsection, the VoI wants to download a large file from the remote server. This requested large file may be first split into multiple pieces and transmitted to different infrastructure points so that each infrastructure point has a different piece of data, which enables cooperation among infrastructure. Data delivered to an infrastructure point may be further split and transmitted to either the VoI or *helpers* when they move into its coverage so that *helpers* have different pieces of data from each other and from the VoI. Data received by the *helpers* will be transmitted to the VoI when they encounter the VoI, which exploits the mobility of vehicles and V2V communications to achieves the vehicular cooperations. Since vehicles in the same direction move at the same constant speed, the inter-vehicle distances in the same lane remain the same at any time instant. This follows that vehicles in the same direction of the VoI can only offer limited help to the VoI because only vehicles within the

coverage of the VoI can offer help. Therefore, in this paper, we only consider vehicles in the opposite direction of the VoI that will receive data from infrastructure and can transmit the received data to the VoI as *helpers*.

To present the cooperative communication strategy, it suffices to consider two consecutive infrastructure points along the travel direction of the VoI collaborating to deliver data. Denote the nearest infrastructure point along the travel direction of the VoI by  $I_1$  and the second nearest one by  $I_2$ . When the VoI is in the coverage of  $I_1$ , it receives data directly from  $I_1$ . In the meantime, the *helpers* may also receive different pieces of data from  $I_2$  when they move through the coverage of  $I_2$ . When the VoI moves outside the coverage of  $I_1$ , it may continue to receive data from *helpers* using V2V communications. Of course, when the VoI moves along its direction, the two infrastructure points participating in the cooperative communication are also updated. In this way, V2I communications between the VoI and infrastructure, between *helpers* and their respective infrastructure points, V2V communications between the VoI and *helpers*, as well as vehicular mobility are coherently combined to maximize the throughput of the VoI. In our considered network scenario, V2I communications by both the VoI and *helpers* are essential to retrieve data from the infrastructure. V2V communications only help to assist the VoI to retrieve more data from the Internet and deliver the data retrieved by *helpers* to the VoI. V2V communications can not increase the net amount of data in the network. Furthermore, we consider that some practical issues like out of sequence data delivery can be handled by techniques such as network coding (e.g., our recent paper [34]) so that we can focus on the main theme of the paper without the need for considering their impacts.

### D. Problem Formation

Now we give a formal definition of the throughput studied in this paper. Consider an arbitrarily chosen time interval  $[0, t]$  and denote the amount of data received by the VoI as  $D(t)$ , which includes data received from both infrastructure and *helpers*. In this paper, we are interested in finding the long-term achievable throughput of the VoI, using our cooperative communication strategy, where the long-term throughput, denoted by  $\eta$ , is formally defined as follows:

$$\eta = \lim_{t \rightarrow \infty} \frac{D(t)}{t}. \quad (1)$$

Without loss of generality, we assume that the VoI travels at speed  $v_1$ , and the *helpers* travel at speed  $v_2$  and have vehicular density  $\rho_2$ . We define the time interval starting from the time instant when the VoI enters into the coverage of one infrastructure point to the time instant when the VoI enters into the coverage of the next infrastructure point as *one cycle*. By using *cycles* as the basic blocks, the entire data receiving process of the VoI can be modeled by a renewal reward process [35]. Each *cycle* in the renewal reward process consists of one V2I communication process, followed by a V2V communication process, and the reward is the amount of

data received by the VoI during each *cycle*. This follows that the throughput can be calculated as follows:

$$\eta = \lim_{t \rightarrow \infty} \frac{D(t)}{t} = \frac{E[D_I] + E[D_V]}{E[T]}, \quad (2)$$

where  $E[D_I]$  and  $E[D_V]$  are respectively the expected amount of data received by the VoI during the V2I communication process and the V2V communication process in one *cycle*, and  $E[T]$  is the expected time of one *cycle*, which can be calculated as  $E[T] = \frac{d}{v_1}$ . Since when the VoI is covered by an infrastructure point, it will only use V2I communication,  $E[D_I]$  can be readily obtained as follows:

$$E[D_I] = \frac{2r_I w_I}{v_1}. \quad (3)$$

Using (2) and (3), the problem of calculating the achievable throughput by the VoI can be transformed into the problem of calculating the expected amount of data that can be received by the VoI from V2V communications in one *cycle*. Without loss of generality, we call the two infrastructure points  $I_1$  and  $I_2$  respectively as defined earlier. Because of the unicast transmission model we adopt, during V2V communications between the VoI and *helpers*, the amount of data two adjacent *helpers* can deliver to the VoI become correlated when their inter-vehicle distance is smaller than  $2r_0$  and is further limited by the amount of data each *helper* receives from  $I_2$ , which can also be correlated because during *helpers'* V2I communications, the amount of data received by adjacent *helpers* become correlated when their inter-vehicle distance is smaller than  $2r_I$ . This complicated correlation structure is quite intricate for statistical analysis. In this paper, we handle the challenge by formulating the V2V data delivering process in one *cycle* as a constrained optimization problem, with the goal of obtaining the maximum amount of data received by the VoI from *helpers* and finding the corresponding scheduling scheme, which includes V2I transmission scheme for *helpers* and V2V transmission scheme, to reach this maximum value. In the following, we will show the formation of the constrained optimization problem.

Denote by  $n$  the number of *helpers* encountered by the VoI during a *cycle* and  $n$  is a Poissonly distributed random integer. Denote by  $V_1$  the first *helper* encountered by the VoI when the VoI moves outside the coverage of  $I_1$ , by  $V_2$  the second *helper*, and so on. Denote the distance between two consecutive *helpers*  $V_i$  and  $V_{i+1}$  by  $l_i, i = 1, \dots, n-1$ . See Fig. 2 for an illustration. Furthermore, denote by  $D_i$  the amount of data received by *helper*  $V_i$  from  $I_2$ , and by  $Y_i$  the amount of data delivered by  $V_i$  to the VoI. We first consider the situation that  $n$  is a fixed integer and  $l_i, i = 1, \dots, n-1$  are known values, i.e., corresponding to a specific instance of these random values, and then extend to consider the more general situation that  $n$  and  $l_i, i = 1, \dots, n-1$  are random values.

Without considering the boundary case, caused by *helpers* located near the borders of the coverage area of infrastructure points, the problem of finding the maximum amount of data received by the VoI from V2V communications in one *cycle*, given  $n$  and  $l_i, i = 1, \dots, n-1$ , can be formulated as the

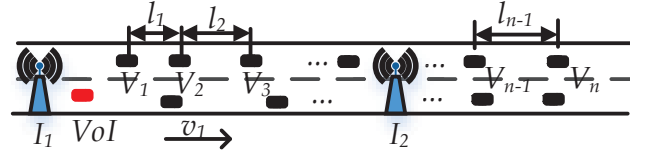


Fig. 2. An illustration of *helpers* encountered by the VoI during one V2V communication *cycle* and their interval distance.

following optimization problem. We will show later in the simulation that the boundary case has negligible impact on the achievable throughput as we are focusing on the long-term throughput. The optimization is taken over the set of all possible scheduling schemes. Without causing any confusion, we drop the notation for the set of all possible scheduling schemes to have a simpler expression.

$$\max \sum_{i=1}^n Y_i \quad (4)$$

$$s.t. 0 \leq D_i \leq \frac{2r_I}{v_2} w_I, \quad i = 1, 2, \dots, n \quad (5)$$

$$\sum_{i=k_1}^{k_2} D_i \leq \frac{\sum_{i=k_1}^{k_2-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I, \quad 1 \leq k_1 \leq k_2 \leq n \quad (6)$$

$$0 \leq Y_i \leq \frac{2r_0}{v_1 + v_2} w_V, \quad i = 1, 2, \dots, n \quad (7)$$

$$Y_i \leq D_i, \quad i = 1, 2, \dots, n \quad (8)$$

$$\sum_{i=k_1}^{k_2} Y_i \leq \frac{\sum_{i=k_1}^{k_2-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V, \quad 1 \leq k_1 \leq k_2 \leq n \quad (9)$$

In the above optimization problem,  $\sum_{i=1}^n Y_i$  is the total amount of data received by the VoI from *helpers* during one *cycle*. Constraint (5) gives the maximum and the minimum amount of data received by each *helper* from infrastructure point  $I_2$ . Constraint (6) gives an upper bound on the amount of data that any  $k_2 - k_1 + 1$  consecutive *helpers* can receive from  $I_2$ , where the term  $\frac{\sum_{i=k_1}^{k_2-1} \min\{l_i, 2r_I\} + 2r_I}{v_2}$  gives the total amount of time these  $k_2 - k_1 + 1$  consecutive *helpers* can receive data from  $I_2$ . Particularly, due to the randomness of vehicle distributions, it may happen that there exists a void region of larger than  $2r_I$ , which has no vehicle (*helper*). When the void region occurs, *helpers* may not be able to receive data continuously from  $I_2$ . Therefore, all two constraints (5) and (6) must be considered to completely describe the V2I communication between  $I_2$  and *helpers*. Constraint (6) also captures the correlation that may occur during the data receiving process of adjacent *helpers*, which has been explained earlier. Similarly, constraint (7) gives the maximum and minimum amount of data that can be received by the VoI from each *helper*. Constraint (8) implies that the amount of data each *helper* can deliver to the VoI cannot exceed the data it receives from  $I_2$ . Constraint (9) gives the upper bound of the amount of data the VoI can receive from any  $k_2 - k_1 + 1$  consecutive *helpers*.

#### IV. THEORETICAL ANALYSIS OF V2V COMMUNICATION PROCESS AND ACHIEVABLE THROUGHPUT

The data received by the VoI, whether through V2I communications directly from the infrastructure or through V2V communications from *helpers*, eventually comes from the infrastructure. Intuitively, as we increase the data rate of V2I communications,  $w_I$ , from a very small value while keeping other parameters constant, initially the throughput will be limited by the data rate of V2I communications. We call this regime the *Infrastructure-Limited Regime*. As we further increase the value of  $w_I$ , we will reach a *Transitional Regime* where both the data rate of V2I communications and the data rate of V2V communications play major roles in determining the throughput of the VoI. If we increase the value of  $w_I$  further to a very large value, V2I communications will no longer be a bottleneck in determining the throughput of the VoI. Instead, the data rate of V2V communication,  $w_V$ , becomes the determining factor of the VoI throughput. We call this regime the *V2V-Limited Regime*.

It is evident from the optimization problem (4) that the amount of data received from V2V communications by the VoI given fixed  $n$  and  $l_i, i = 1, \dots, n$ ,  $\sum_{i=1}^n Y_i$ , satisfies:

$$\sum_{i=1}^n Y_i \leq \min\{D_{Vu1}, D_{Vu2}\} \quad (10)$$

where  $D_{Vu1} = \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I$  comes from a combination of constraints (6) and (8), representing the maximum amount of data all *helpers* can receive from infrastructure; and  $D_{Vu2} = \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V$  comes from constraint (9), representing the maximum amount of data all *helpers* can deliver to the VoI through V2V communications without considering the limitation of the amount of data they receive. When  $0 < w_I \leq \frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$ , we have  $D_{Vu1} \leq D_{Vu2}$ , which implies that the amount of data the VoI can receive from *helpers* is limited by the amount of data *helpers* can receive through their V2I communications, thus limited by  $w_I$ . Thus, we define the *Infrastructure-Limited Regime*:  $0 < w_I \leq \frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$ . Similarly, when  $w_I \geq \frac{w_V v_2}{v_1 + v_2}$ , we have  $D_{Vu1} > D_{Vu2}$ , which implies that the amount of data the VoI can receive from *helpers* is limited by the amount of data the *helpers* can deliver through V2V communications, thus limited by  $w_V$ . Therefore, we define the *V2V-Limited Regime*:  $w_I \geq \frac{w_V v_2}{v_1 + v_2}$ . The rest of the region forms the *Transitional Regime*:  $\frac{r_0 w_V v_2}{r_I (v_1 + v_2)} < w_I < \frac{w_V v_2}{v_1 + v_2}$ .

In the following subsections, we analyze the achievable throughput by the VoI under each regime separately.

##### A. Infrastructure-Limited Regime: $0 < w_I \leq \frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$

In this subsection, we first analyze the maximum amount of data that can be received from *helpers* by the VoI in one cycle by solving the optimization problem (4), and find the corresponding scheduling scheme to achieve this maximum solution given fixed  $n$  and  $l_i, i = 1, \dots, n-1$ . Then, we extend to consider that  $n$  and  $l_i, i = 1, \dots, n-1$  are random values, corresponding to Poisson distribution of vehicles, and analyze the achievable throughput under the obtained optimal scheduling scheme.

1) *An analysis of the V2V communication process* : The following theorem summarizes the major result of this subsection.

**Theorem 1.** *In the Infrastructure-Limited regime, given fixed  $n$  and  $l_i, i = 1, 2, \dots, n-1$ , the maximum amount of data the VoI can receive from all  $n$  helpers in one cycle is given by*

$$\left( \sum_{i=1}^n Y_i \right)_1^* = \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I \quad (11)$$

where  $(\sum_{i=1}^n Y_i)_1^*$  is the respective  $\sum_{i=1}^n Y_i$  associated with its optimum value and we use the subscript 1 to mark the Infrastructure-Limited regime and superscript \* to mark the optimum value.

Furthermore, there exists a V2I transmission scheme for *helpers* and a V2V transmission scheme to reach the above maximum amount of received data for the VoI, satisfying:

$$\begin{cases} Y_{1i}^* = D_{1i}^* = \frac{\min\{l_i, 2r_I\}}{v_2} w_I, & i = 1, 2, \dots, n-1 \\ Y_{1n}^* = D_{1n}^* = \frac{2r_I}{v_2} w_I \end{cases} \quad (12)$$

where  $D_{1i}^*$  and  $Y_{1i}^*, i = 1, \dots, n$  are the respective  $D_i$  and  $Y_i, i = 1, \dots, n$  associated with the optimum solution.

*Proof:* In the Infrastructure-Limited regime, with conditions  $w_I \leq \frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$  and  $r_I > r_0$ , we have  $\frac{2r_I}{v_2} w_I \leq \frac{2r_0}{v_1 + v_2} w_V$  and  $\frac{\min\{l_i, 2r_I\}}{v_2} w_I \leq \frac{\min\{l_i, 2r_0\}}{v_1 + v_2} w_V, i = 1, \dots, n-1$ . It follows that:

$$\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I \leq \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V \quad (13)$$

The above equation implies that in the optimization problem (4), constraint (7) is redundant and constraint (9) can be replaced with a tighter constraint after merging the two constraints (6) and (8). The new constraints for the optimization problem (4) under Infrastructure-Limited regime are shown as follows:

$$0 \leq Y_i \leq D_i \leq \frac{2r_I}{v_2} w_I, \quad i = 1, 2, \dots, n \quad (14)$$

$$\sum_{i=k_1}^{k_2} Y_i \leq \sum_{i=k_1}^{k_2} D_i \leq \frac{\sum_{i=k_1}^{k_2-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I, \quad 1 \leq k_1 \leq k_2 \leq n \quad (15)$$

Constraint (15) shows that an upper bound of  $\sum_{i=1}^n Y_i$  is given by  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I$ . In the following, we will show that this upper bound is exactly the optimum solution of  $\sum_{i=1}^n Y_i$  in the Infrastructure-Limited regime and can be reached under some scheduling scheme.

Noting that the upper bound of  $\sum_{i=1}^n Y_i$ ,  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I$ , is the sum of  $n$  separate components, with each component smaller than or equal to  $\frac{2r_I}{v_2} w_I$ . Therefore, when each  $Y_i$  is equal to  $D_i$ , and is further equal to the corresponding component forming  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I$ , i.e., when each  $Y_i$  and each  $D_i, i = 1, \dots, n$  are given by (12), the value of  $\sum_{i=1}^n Y_i$  will reach its upper bound  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I$  while



satisfying the constraints in optimization problem (4). This leads to the expression of (11).

It remains to demonstrate that there exists a scheduling scheme to reach this optimum solution specified in (11). To this end, we show that (12) readily leads to the design of an optimal transmission scheme. Specifically, a scheduling scheme which schedules both V2I and V2V transmissions on a first-in-first-out (FIFO) basis can achieve the optimum solution (11). We acknowledge that the optimum scheduling algorithm that achieves the optimum solution (11) may not be unique. When other performance metric is considered, e.g., delay, the earliest deadline first scheme may have better delay performance while achieving the same throughput. Particularly, in the FIFO scheduling scheme, each *helper* starts its V2I communication once it enters the coverage of infrastructure point  $I_2$  and there is no other *helper* preceding it communicating with the infrastructure point  $I_2$ , and stops when the *helper* leaves the coverage of the infrastructure point  $I_2$ , which lead to that each *helper* will receive an amount of data shown as each  $D_{1i}^*$ ,  $i = 1, 2, \dots, n$  in Eq. (12). Similarly, for V2V communications, the VoI receives data from each *helper* one by one when there exists at least one *helper* within its coverage on a FIFO basis. Once a *helper* starts to deliver its data to the VoI, it will stop until it has transmitted all its data to the VoI or when it leaves the coverage of the VoI, which lead to that the data the VoI receives from each *helper* is shown as each  $Y_{1i}^*$ ,  $i = 1, 2, \dots, n$  in Eq. (12). Noting that Eq. (12) leads to the optimum solution (11), therefore, it can be readily established that the aforementioned scheduling scheme achieves the maximum amount of the received data for the VoI specified in (11). ■

*Remark 1.* Note that (11) and the corresponding scheduling scheme that satisfies (12) are valid for any value of  $n$  and the corresponding  $l_i, i = 1, \dots, n - 1$ .

2) *Throughput calculation:* On the basis of Theorem 1, we now analyze the achievable throughput by the VoI considering that both  $n$  and the corresponding  $l_i, i = 1, \dots, n - 1$  are random values. A brute force approach of computing the achievable throughput will first consider that  $n$  is a Poisson random variable, then conditioned on each value of  $n$  (noting that conditional on a specific instance of  $n$ , *helpers* become uniformly distributed and hence  $l_i, i = 1, \dots, n - 1$ , become correlated), evaluate the joint distribution of the random variables  $\min\{l_i, 2r_I\}$ ,  $i = 1, \dots, n - 1$ , and finally transform the conditional value into unconditional one using the total probability theorem and the Poisson distribution of  $n$ . This will result in a very complicated analysis. In the following, we use simpler techniques by resorting to the concept of *clusters*, defined shortly later, to analyze the achievable throughput.

We designate the time instant when the VoI leaves the coverage of  $I_1$  as  $t = 0$  and define its moving direction as the positive (right) direction of the coordinate system. Furthermore, we define the point to the right of  $I_1$  and at a distance  $r_I - r_0$  to  $I_1$  as the origin of the coordinate system. It follows from the above that the time instant when the VoI enters into  $I_2$ 's coverage will be  $t_1 = \frac{d-2r_I}{v_1}$ . Noting that the relative speed of the VoI to the *helpers* traveling in the opposite

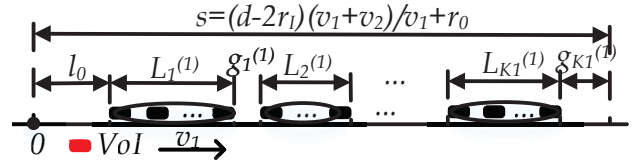


Fig. 3. An illustration of *clusters* formed by the *helpers*. Each *cluster* has length  $L_j^{(1)}, j = 1, \dots, K_1$  and each gap between two consecutive *clusters* has length  $g_j^{(1)}, j = 1, \dots, K_1$ .

direction is  $v_1 + v_2$ , therefore the relative distance traveled by the VoI, relative to the *helpers* in the opposite direction which all travel at the same constant speed of  $v_2$  during  $[0, t_1]$ , is given by  $\frac{(d-2r_I)}{v_1}(v_1 + v_2)$ . The random number of *helpers* encountered by the VoI, who may deliver data to the VoI, during  $[0, t_1]$ , is determined by the parameter  $s$ :

$$s = \frac{(d - 2r_I)(v_1 + v_2)}{v_1} + r_0, \quad (16)$$

where the  $r_0$  term is due to the consideration that when the VoI exits the coverage of  $I_1$  and is located at coordinate  $r_0$  (and at time instant  $t = 0$ ), the *helper(s)* to the left of the VoI and within a distance  $r_0$  to the VoI may possibly deliver data to the VoI too. Thus, all *helpers* in the opposite direction that the VoI may encounter during its V2V communication process in one *cycle* are within road segment  $[0, s]$ .

As explained in the beginning of this subsection, we use the concept of *clusters* to simplify our analysis. A *cluster* is defined as a maximal set of *helpers* located within road segment  $[0, s]$  and the distance between any two adjacent *helpers* is smaller than or equal to  $2r_I$ . Forming *clusters* in this way allows us to remove the complexity associated with the computation of the joint distribution of  $\min\{l_i, 2r_I\}, i = 1, \dots, n-1$  because within each *cluster*, we have  $\min\{l_i, 2r_I\} = l_i, i = 1, 2, \dots$ . For each *cluster*, we only need to focus on the length of the *cluster* rather than the individual inter-vehicle distances. There may be multiple *clusters* within road segment  $[0, s]$  and a *cluster* may contain single vehicle only. Denote the coordinate of the first *helper* that can transmit data to the VoI since  $t = 0$  by  $l_0$ . Due to the memoryless property of the exponential distribution of inter-vehicle distances,  $l_0$  has the same exponential distribution as other  $l_i, i = 1, \dots, n - 1$  and the starting position of a vehicle in a *cluster* does not affect the distribution of the length of the *cluster*. Denote by  $K_1$  the random non-negative integer representing the number of *clusters* the VoI will encounter in one *cycle*. Furthermore, denote by  $L_j^{(1)}, j = 1, \dots, K_1$ , the length of each *clusters*, which are identically and independently distributed (i.i.d), and by  $g_j^{(1)}, j = 1, \dots, K_1$  the length of each gap between two adjacent *clusters*, which are also i.i.d. See Fig. 3 for an illustration.

Noting that (11) is also valid for any subset of *helpers* within road segment  $[0, s]$  adopting the scheduling scheme described in the proof of Theorem 1, therefore the amount of data each *cluster* of *helpers* delivers to the VoI, denoted by  $R_j^{(1)}, j = 1, \dots, K_1$ , can be obtained as follows (recall that the analysis is



conducted for the Infrastructure-Limited regime):

$$R_j^{(1)} = \frac{L_j^{(1)} + 2r_I}{v_2} w_I, j = 1, \dots, K_1 \quad (17)$$

It follows that the amount of data received by the VoI from *helpers* in one *cycle*, denoted by  $D_{V1}$ , can be readily calculated by summing up the amount of data received by the VoI from each *cluster* of *helpers*:

$$D_{V1} = \sum_{j=1}^{K_1} R_j^{(1)} = \sum_{j=1}^{K_1} \frac{L_j^{(1)} + 2r_I}{v_2} w_I. \quad (18)$$

Noting that in (18), both the number of *clusters* in road segment  $[0, s]$ ,  $K_1$ , and the length of each *cluster*,  $L_j^{(1)}$ , are random variables, and they are not independent. If we *approximately* consider they are independent with each other, then from (18), the expected amount of data received by the VoI from V2V communications in one *cycle*,  $E[D_{V1}]$ , can be calculated as follows :

$$E[D_{V1}] = E[K_1] \cdot \frac{E[L_j^{(1)}] + 2r_I}{v_2} w_I, \quad (19)$$

where  $E[K_1]$  is the expected number of *clusters* in the road segment  $[0, s]$ . The accuracy of this *approximation* is verified by simulation, see Fig. 4 below. It shows that the *approximation* will marginally increase the result of  $E[D_{V1}]$ , which in turn, will leads to a marginally increase in the achievable throughput in this regime.

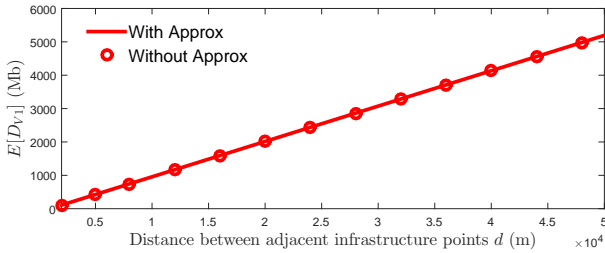


Fig. 4. A comparison between the result of  $E[D_{V1}]$  with and without the approximation.

As  $L_j^{(1)}, j = 1, \dots, K_1$  and  $g_j^{(1)}, j = 1, \dots, K_1$  are both i.i.d and each  $L_j^{(1)}$  and  $g_j^{(1)}$  are also mutually independent, then according to the Generalized Wald's equality [35, Theorem 4.5.2], when  $s \gg E[L_j^{(1)}] + E[g_j^{(1)}]$ ,  $E[K_1]$  can be approximately calculated as follows:

$$E[K_1] = \frac{s - E[l_0]}{E[L_j^{(1)}] + E[g_j^{(1)}]}. \quad (20)$$

By putting (20) into (19), we have

$$E[D_{V1}] = \frac{s - E[l_0]}{E[L_j^{(1)}] + E[g_j^{(1)}]} \times \frac{E[L_j^{(1)}] + 2r_I}{v_2} w_I, \quad (21)$$

where the values of  $E[L_j^{(1)}]$  and  $E[g_j^{(1)}]$  have been given by [22]:

$$E[L_j^{(1)}] = (e^{2\rho_2 r_I} - 1) \left( \frac{1}{\rho_2} - \frac{2r_I e^{-2\rho_2 r_I}}{1 - e^{-2\rho_2 r_I}} \right), \quad (22)$$

and

$$E[g_j^{(1)}] = 2r_I + \frac{1}{\rho_2}. \quad (23)$$

As mentioned before, due to the memoryless property of exponential distribution,  $l_0$  has the same distribution as  $l_i$  [36], i.e., we have:

$$E[l_0] = \frac{1}{\rho_2}. \quad (24)$$

Combining (21)-(24), we can obtain:

$$E[D_{V1}] = \frac{\left[ \frac{(d-2r_I)(v_1+v_2)}{v_1} + r_0 - \frac{1}{\rho_2} \right] (1 - e^{-2\rho_2 r_I}) w_I}{v_2} \quad (25)$$

By plugging equations (3) and (25) into (2), we have the achievable throughput in Infrastructure-Limited regime, denoted by  $\eta_1$ , as follows:

$$\eta_1 = \frac{2r_I w_I + c_1}{d}, \quad (26)$$

where

$$c_1 = \frac{\left[ (d-2r_I)(v_1+v_2) + r_0 v_1 - \frac{v_1}{\rho_2} \right] (1 - e^{-2\rho_2 r_I}) w_I}{v_2}.$$

#### B. V2V-Limited Regime: $w_I \geq \frac{w_V v_2}{v_1 + v_2}$

Now we analyze the achievable throughput in the V2V-Limited regime. Similar as subsection IV-A, in this subsection, we first analyze the maximum amount of data that can be received from *helpers* by the VoI in one *cycle* by solving the optimization problem (4), and then find the corresponding scheduling scheme to achieve this maximum solution given fixed  $n$  and  $l_i, i = 1, \dots, n-1$ . Finally, we extend to consider that  $n$  and  $l_i, i = 1, \dots, n-1$  are random values, corresponding to Poisson distribution of vehicles, and analyze the achievable throughput under the proposed scheduling scheme.

1) *An analysis of the V2V communication process:* The following theorem summarizes the main result of this subsection.

**Theorem 2.** *In the V2V-Limited regime, given fixed  $n$  and  $l_i, i = 1, \dots, n-1$ , the maximum amount of data the VoI can receive from all  $n$  helpers in one cycle is given by*

$$\left( \sum_{i=1}^n Y_i \right)_2^* = \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V, \quad (27)$$

where  $(\sum_{i=1}^n Y_i)_2^*$  is the respective  $\sum_{i=1}^n Y_i$  associated with its optimum value and we use the subscript 2 to mark the V2V-Limited regime.

Furthermore, there exists a V2I transmission scheme for *helpers* and a V2V transmission scheme to reach the above maximum amount of received data for the VoI, satisfying:

$$\begin{cases} D_{2i}^* = \frac{\min\{l_i, 2r_I\}}{v_2} w_I, i = 1, 2, \dots, n-1 \\ D_{2n}^* = \frac{2r_I}{v_2} w_I \\ Y_{2i}^* = \frac{\min\{l_i, 2r_0\}}{v_1 + v_2} w_V, i = 1, 2, \dots, n-1 \\ Y_{2n}^* = \frac{2r_0}{v_1 + v_2} w_V \end{cases} \quad (28)$$

where  $D_{2i}^*$  and  $Y_{2i}^*, i = 1, \dots, n$  are the respective  $D_i$  and  $Y_i, i = 1, \dots, n$  associated with the optimum solution.

*Proof:* In the V2V-Limited regime, with conditions  $w_I \geq \frac{w_V v_2}{v_1 + v_2}$  and  $r_I > r_0$ , we have  $\frac{2r_0}{v_1 + v_2} w_V < \frac{2r_I}{v_2} w_I$  and  $\frac{\min\{l_i, 2r_0\}}{v_1 + v_2} w_V \leq \frac{\min\{l_i, 2r_I\}}{v_2} w_I, i = 1, \dots, n-1$ . It follows that:

$$\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V < \frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_I\} + 2r_I}{v_2} w_I, \quad (29)$$

Then from constraints (6), (8), (9) and inequality (29), we can conclude that the value of  $\sum_{i=1}^n Y_i$  in optimization problem (4) is upper bound by  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V$ . In the following, we will show that this upper bound is exactly the optimum solution of  $\sum_{i=1}^n Y_i$  in the V2V-Limited regime and can be reached under some scheduling scheme.

Noting that  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V$  is the sum of  $n$  separate components, with each component is not larger than  $\frac{2r_0}{v_1 + v_2} w_V$ . Therefore, when each  $Y_i$  is equal to the corresponding part of the  $n$  component forming the upper bound of  $\sum_{i=1}^n Y_i$ ,  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V$ , and each  $D_i$  is equal to the corresponding part of the  $n$  component forming the upper bound of  $\sum_{i=1}^n D_i$  shown in (6), i.e., when each  $Y_i$  and  $D_i, i = 1, \dots, n$  are given by (28), the value of  $\sum_{i=1}^n Y_i$  will reach its upper bound  $\frac{\sum_{i=1}^{n-1} \min\{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V$  while satisfying all constraints in optimization problem (4). This leads to the expression of (27).

Now we show that there exists a scheduling scheme to reach the maximum solution specified in (27). To this end, we show that (28) readily leads to the design of the scheduling scheme. Particularly, the scheduling scheme schedules both *helpers'* V2I communication and V2V communications on a FIFO basis. Specifically, the V2I transmission scheme for *helpers* is the same as that for the Infrastructure-Limited regime, which lead to that each *helper* will receive an amount of data shown as each  $D_{2i}^*, i = 1, 2, \dots, n$  in Eq. (28). For V2V communications, the VoI starts to receive data from a *helper* once it enters this *helper's* coverage and has retrieved all data from the previous *helper* or has left the previous *helper's* coverage, and stops when the VoI leaves the coverage of the current *helper* or has retrieved all data of the current *helper*, which lead to that the data the VoI receives from each *helper* is shown as each  $Y_{2i}^*, i = 1, 2, \dots, n$  in Eq. (28). Noting that Eq. (28) leads to the optimum solution (27), therefore, it can be readily established that the aforementioned scheduling scheme achieves the maximum amount of the received data for the VoI specified in (27). This completes the proof. ■

Similarly as Theorem 1, Theorem 2 is also valid for any value of  $n$  and the corresponding  $l_i, i = 1, \dots, n-1$ .

2) *Throughput calculation:* On the basis of Theorem 2, we now analyze the achievable throughput by the VoI considering that both  $n$  and the corresponding  $l_i, i = 1, \dots, n-1$  are random values. In the V2V-Limited regime, we define a *cluster* to be a maximal set of *helpers* located within road segment  $[0, s]$  and the distance between any two adjacent *helpers* is smaller than or equal to  $2r_0$ . The reason that we define the *clusters* differently from that in the Infrastructure-Limited regime is that in this regime, it is the correlation in the V2V communication process (and the associated difficulty in determining the joint distribution of  $\min\{l_i, 2r_0\}, i = 1, \dots, n-1$ ) that plays a

dominating effect on determining the achievable throughput. By defining *clusters* in the above way, within each *cluster*, we have  $\min\{l_i, 2r_0\} = l_i, i = 1, 2, \dots$ .

In the V2V-Limited regime, the amount of data each *cluster* of *helpers* delivers to the VoI, denoted by  $R_j^{(2)}$ , can be calculated as follows:

$$R_j^{(2)} = \frac{L_j^{(2)} + 2r_0}{v_1 + v_2} w_V, j = 1, \dots, K_2 \quad (30)$$

where  $L_j^{(2)}, j = 1, \dots, K_2$  is the random variable representing the length of the  $j$ -th *cluster*, and  $K_2$  is the random integer representing the number of *clusters* the VoI will encounter in one *cycle*.

Utilizing the same *approximation* method as that used to calculate  $E[D_{V1}]$  in the Infrastructure-Limited regime, i.e., *approximately* consider that  $K_2$  and  $L_j^{(2)}$  are independent in this regime, the expected amount of data received by the VoI from *helpers* in one *cycle* in the V2V-Limited regime can be obtained as follows:

$$E[D_{V2}] = \frac{s - E[l_0]}{E[L_j^{(2)}] + E[g_j^{(2)}]} \times \frac{E[L_j^{(2)}] + 2r_0}{v_1 + v_2} w_V, \quad (31)$$

where  $E[L_i^{(2)}]$  and  $E[g_i^{(2)}]$  are given by:

$$E[L_j^{(2)}] = (e^{2\rho_2 r_0} - 1) \left( \frac{1}{\rho_2} - \frac{2r_0 e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_0}} \right), \quad (32)$$

and

$$E[g_j^{(2)}] = 2r_0 + \frac{1}{\rho_2}. \quad (33)$$

Combing (24) and (31)-(33), and simplifying it, we have:

$$E[D_{V2}] = \frac{\left[ \frac{(d-2r_I)(v_1+v_2)}{v_1} + r_0 - \frac{1}{\rho_2} \right] (1 - e^{-2\rho_2 r_0}) w_V}{v_1 + v_2}. \quad (34)$$

By plugging (3), (34) into (2), the achievable throughput in the V2V-Limited regime, denoted by  $\eta_2$ , can be obtained as follows:

$$\eta_2 = \frac{2r_I w_I + c_2}{d}, \quad (35)$$

where

$$c_2 = \frac{\left[ (d-2r_I)(v_1+v_2) + r_0 v_1 - \frac{v_1}{\rho_2} \right] (1 - e^{-2\rho_2 r_0}) w_V}{v_1 + v_2}.$$

C. *Transitional Regime:*  $\frac{r_0 w_V v_2}{r_I (v_1 + v_2)} < w_I < \frac{w_V v_2}{v_1 + v_2}$

Now we analyze the achievable throughput in the transitional regime where the analysis is more intricate than that for the Infrastructure-Limited and the V2V-Limited regime. Particularly, in the transitional regime, both V2V communications and *helpers'* V2I communications contribute to determining the achievable throughput of the VoI. Therefore both the correlation in the amount of data received by adjacent *helpers* from infrastructure and in the amount of data received by the VoI from adjacent *helpers*, as explained in Section III-D, need to be considered. This makes finding the optimum solution for the optimization problem (4) more challenging. Therefore, in this subsection, instead of analyzing the exact achievable

throughput, we analyze its upper and lower bound. In the following, we will analyze the upper and the lower bound of the achievable throughput separately.

1) *Upper bound of the achievable throughput:* As shown in (10), an upper bound of  $\sum_{i=1}^n Y_i$  is given by:

$$\sum_{i=1}^n Y_i \leq \min\{D_{Vu1}, D_{Vu2}\} \quad (36)$$

That is, the upper bound of data amount received by the VoI from *helpers* is determined by the smaller value of the amount of data received by the *helpers* from their V2I communications,  $D_{Vu1}$ , and the amount of data *helpers* can deliver to the VoI in V2V communications (without considering the limitation of the amount of data they receive),  $D_{Vu2}$ . It is shown in Theorem 1 and Theorem 2 that  $D_{Vu1}$  and  $D_{Vu2}$  are exactly the maximum amount of data the VoI can receive from *helpers* in the Infrastructure-Limited regime and V2V-Limited regime respectively. Therefore, according to the throughput calculation analysis given in subsection IV-A and IV-B, when  $\sum_{i=1}^n Y_i$  is upper bounded by  $D_{Vu1}$  (or  $D_{Vu2}$ ), the corresponding achievable throughput of the VoI will be upper bounded by the achievable throughput in the Infrastructure-Limited regime,  $\eta_1$ , (or the achievable throughput in the V2V-Limited regime,  $\eta_2$ ). It follows that an upper bound of the achievable throughput by the VoI in the transitional regime, denoted by  $\eta_{3u}$ , is given by:

$$\eta_{3u} = \min\{\eta_1, \eta_2\} = \begin{cases} \eta_1, & \eta_1 \leq \eta_2 \\ \eta_2, & \eta_1 > \eta_2 \end{cases} \quad (37)$$

Putting (26) and (35) into (37) and simplifying it, we have:

$$\eta_{3u} = \begin{cases} \frac{2r_I w_I + c_1}{d}, \frac{r_0 w_V v_2}{r_I (v_1 + v_2)} < w_I \leq \frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2} \\ \frac{2r_I w_I + c_2}{d}, \frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2} < w_I < \frac{w_V v_2}{v_1 + v_2} \end{cases} \quad (38)$$

with  $c_1$  and  $c_2$  have been given in the earlier analysis.

*Remark 2.* Equation (38) shows that  $\frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2}$  is a transition point for the value of  $w_I$  to determine the upper bound of achievable throughput in the transitional regime, whose value depends on the *helpers'* density  $\rho_2$ . Specifically, when  $\rho_2 \rightarrow 0$ ,  $\frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2} \rightarrow \frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$ ; when  $\rho_2$  increases, the gap between  $\frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2}$  and  $\frac{r_0 w_V v_2}{r_I (v_1 + v_2)}$  becomes larger and the gap between  $\frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2}$  and  $\frac{w_V v_2}{v_1 + v_2}$  becomes smaller; and when  $\rho_2 \rightarrow \infty$ ,  $\frac{1 - e^{-2\rho_2 r_0}}{1 - e^{-2\rho_2 r_I}} \times \frac{w_V v_2}{v_1 + v_2} \rightarrow \frac{w_V v_2}{v_1 + v_2}$ .

2) *Lower bound of the achievable throughput:* In this subsection, we first analyze the lower bound of the maximum amount of data that can be received from *helpers* by the VoI and the corresponding scheduling scheme to achieve this lower bound given fixed  $n$  and  $l_i, i = 1, \dots, n-1$ . Then we extend to consider that  $n$  and  $l_i, i = 1, \dots, n-1$  are random values, corresponding to Poisson distribution of vehicles, and analyze the lower bound of the achievable throughput.

**Theorem 3.** *In the transitional regime, given fixed  $n$  and  $l_i, i = 1, \dots, n-1$ , a lower bound of the maximum amount*

*of data the VoI can receive from  $n$  helpers in one cycle is given by*

$$\left( \sum_{i=1}^n Y_i \right)_3^* \geq \sum_{i=1}^{n-1} \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} + \frac{2r_0}{v_1 + v_2} w_V \quad (39)$$

where  $(\sum_{i=1}^n Y_i)_3^*$  is the respective  $\sum_{i=1}^n Y_i$  associated with its optimum value and we use the subscript 3 to mark the transitional regime.

Furthermore, there exists a V2I transmission scheme for *helpers* and a V2V transmission scheme to achieve the above lower bound of the maximum amount of data for the VoI, satisfying:

$$\begin{cases} D_{3i}^* = \frac{\min\{l_i, 2r_I\}}{v_2} w_I, & i = 1, 2, \dots, n-1 \\ D_{3n}^* = \frac{2r_I}{v_2} w_I \\ Y_{3i}^* = \min \left\{ \frac{l_i w_I}{v_2}, \frac{2r_0 w_V}{v_1 + v_2} \right\}, & i = 1, 2, \dots, n-1 \\ Y_{3n}^* = \frac{2r_0}{v_1 + v_2} w_V \end{cases} \quad (40)$$

where  $D_{3i}^*$  and  $Y_{3i}^*, i = 1, \dots, n$  are the respective  $D_i$  and  $Y_i, i = 1, \dots, n$  associated with the optimum solution.

*Proof:* We find the lower bound of the maximum amount of data received by the VoI from  $n$  *helpers* in one cycle by constructing a specific V2I transmission scheme and analyze the corresponding value of  $\sum_{i=1}^n Y_i$  achieved under this scheme. As this value of  $\sum_{i=1}^n Y_i$  is obtained under a specific V2I transmission scheme, it may not be the maximum value for the original optimization problem (4) because of a lack of consideration of all possible V2I transmission schemes for *helpers*, but will form a lower bound of the maximum value of  $\sum_{i=1}^n Y_i$  for the the original optimization problem (4). In the following, we will first construct a specific V2I transmission scheme, and then analyze the optimum amount of data received by the VoI from *helpers* under this specific V2I transmission scheme, as well as finding a corresponding V2V transmission scheme to reach the lower bound specified in the theorem.

It has been described in the proof of Theorem 1 and Theorem 2 that the V2I transmission scheme for *helpers* to reach the corresponding optimum throughput under the Infrastructure-Limited and the V2V-Limited regime are the same, and this V2I transmission scheme satisfies the following equations:

$$\begin{cases} D_i = \frac{\min\{l_i, 2r_I\}}{v_2} w_I, & i = 1, 2, \dots, n-1 \\ D_n = \frac{2r_I}{v_2} w_I. \end{cases} \quad (41)$$

We adopt this same V2I transmission scheme for *helpers* in the transitional regime as well. It follows that the amount of data received by each *helper* from infrastructure,  $D_i, i = 1, \dots, n$ , is given by (41). With condition  $\frac{r_0 w_V v_2}{r_I (v_1 + v_2)} < w_I < \frac{w_V v_2}{v_1 + v_2}$  for the transitional regime, constraints (7) and (8) in the optimization problem (4) can be replaced with a tighter constraint after putting in (41), which are shown as follows:

$$\begin{aligned} Y_i &\leq \min \left\{ D_i, \frac{2r_0}{v_1 + v_2} w_V \right\} \\ &= \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_I}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} \end{aligned}$$

$$= \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\}, i = 1, 2, \dots, n-1 \quad (42)$$

and

$$0 \leq Y_n \leq \frac{2r_0}{v_1 + v_2} w_V \quad (43)$$

From (42) and (43), the upper bound of  $\sum_{i=1}^n Y_i$  is given by  $\sum_{i=1}^{n-1} \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} + \frac{2r_0}{v_1 + v_2} w_V$ . In the following, we will show that this upper bound is exactly the optimum solution of  $\sum_{i=1}^n Y_i$  under the adopted V2I transmission scheme and we can find a corresponding V2V transmission scheme to achieve this upper bound.

With condition  $w_I < \frac{w_V v_2}{v_1 + v_2}$ , we have:

$$\min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} \leq \min \left\{ \frac{l_i}{v_1 + v_2} w_V, \frac{2r_0}{v_1 + v_2} w_V \right\} \quad (44)$$

This follows that:

$$\begin{aligned} \sum_{i=1}^n Y_i &\leq \sum_{i=1}^{n-1} \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} + \frac{2r_0}{v_1 + v_2} w_V \\ &\leq \frac{\sum_{i=1}^{n-1} \min \{l_i, 2r_0\} + 2r_0}{v_1 + v_2} w_V, \end{aligned} \quad (45)$$

which shows that when  $\sum_{i=1}^n Y_i$  is not larger than  $\sum_{i=1}^{n-1} \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} + \frac{2r_0}{v_1 + v_2} w_V$ , the constraint (9) in optimization problem (4) will also be satisfied. Thus, when each  $Y_i, i = 1, \dots, n$ , is equal to its upper bound shown in (42) and (43), and when each  $D_i, i = 1, \dots, n$  is given by (41), i.e., when each  $Y_i$  and  $D_i, i = 1, \dots, n$  are given by (40), the value of  $\sum_{i=1}^n Y_i$  will reach its upper bound  $\sum_{i=1}^{n-1} \min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} + \frac{2r_0}{v_1 + v_2} w_V$  while satisfying all constraints in the optimization problem (4). This leads to (39).

Now we show that there exists a V2V transmission scheme to achieve the lower bound specified in (39). To this end, it can be readily shown from (40) that the same V2V transmission scheme described in the proof of Theorem 1 satisfies (40), therefore can realize the maximum amount of the received data for the VoI specified in (39) under the specific V2I transmission scheme. This completes the proof. ■

On the basis of Theorem 3, we now analyze the lower bound of the achievable throughput by the VoI considering that both  $n$  and the corresponding  $l_i, i = 1, \dots, n-1$  are random values. Similarly to the analysis in subsections IV-A and IV-B, calculating the lower bound of the achievable throughput directly according to (39) is challenging due to complexity associated with analyzing the joint distribution of  $\min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\}, i = 1, \dots, n-1$ . In this regime, we define a *cluster* to be a maximal set of *helpers* located within road segment  $[0, s]$  and the distance between any two adjacent *helpers* is smaller than or equal to  $\frac{2r_0 w_V v_2}{w_I (v_1 + v_2)}$ . It follows that within each *cluster*,  $\min \left\{ \frac{l_i}{v_2} w_I, \frac{2r_0}{v_1 + v_2} w_V \right\} = \frac{l_i}{v_2} w_I, i = 1, 2, \dots$ , therefore removing the above challenge.

Utilizing the same *approximation* method as that used to calculate  $E[D_{V1}]$  and  $E[D_{V2}]$  in the Infrastructure-Limited regime and the V2V-Limited regime respectively, the lower bound of the achievable throughput in transitional regime, denoted by  $\eta_{3l}$ , is obtained as follows:

$$\eta_{3l} = \frac{2r_I w_I + c_3}{d}, \quad (46)$$

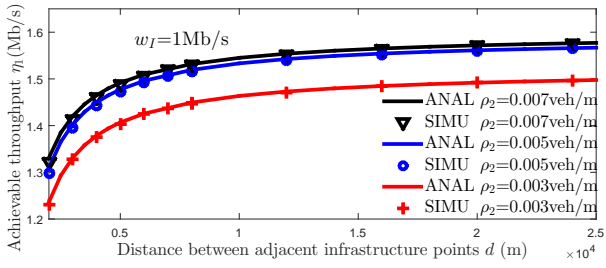
$$\text{where } c_3 = \frac{[(d-2r_I)(v_1+v_2)+r_0v_1-\frac{v_1}{\rho_2}] \left(1 - e^{-\frac{2\rho_2 r_0 w_V v_2}{w_I (v_1+v_2)}}\right) w_I}{v_2}.$$

## V. SIMULATION AND DISCUSSION

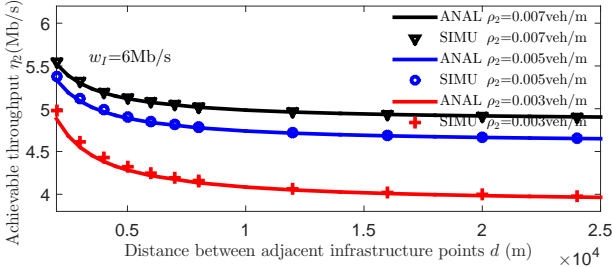
In this section we use simulations conducted in Matlab to verify the accuracy of the analysis and establish the applicability of the theoretical analysis for more general scenarios beyond the ideal assumptions (e.g., constant speed, unit disk model, constant channel condition, etc) used in the analysis. Specifically, 20 infrastructure points are regularly deployed along the highway and the distance between adjacent infrastructure point,  $d$ , is varied from 2km to 50km. The *helpers*' density  $\rho_2$  varies from 0 to 0.1veh/m and the speed of the VoI and *helpers* are  $v_1=15\text{m/s}$  and  $v_2=25\text{m/s}$  respectively. The radio range of infrastructure and vehicles are 500m and 250m (typical radio ranges using DSRC [11]) respectively. The transmission rate of V2V communications is  $w_V=5\text{Mb/s}$  and the transmission rate of V2I communications  $w_I$  varies from 0 to 10Mb/s to allow us to cover all three regimes. Each simulation is repeated 2000 times and the average value is shown in the plot.

Fig. 5 shows a comparison between analytical results and simulation results under each regime. Specifically, Fig. 5(a) and Fig. 5(b) compare the achievable throughput obtained from our analysis and the simulation result, where the simulation is conducted assuming the scheduling scheme described in the proof of Theorem 1 and Theorem 2 respectively. Fig. 5(c) compares the upper and lower bound of the achievable throughput obtained from analysis and the optimum throughput in the simulation. It is shown that in the Infrastructure-Limited and V2V-Limited regime, the analytical results match very well with simulations especially when the distance of two neighboring infrastructure,  $d$ , is large. This confirms that the approximations used in the earlier analysis to obtain the analytical results have negligible impact on the accuracy of the analytical results. In the transitional regime, there is a small gap between the simulated optimum throughput and its upper and lower bound we obtained, e.g, when  $\rho_2=0.004\text{veh/m}$  and  $d=8\text{km}$  in this case, the difference between the optimum throughput from the simulation and its upper (or lower) bound is only around 0.8% (or 5%), and the gap decreases with the increase of *helpers*' density. This shows that even though the upper bound in the transitional regime is not achievable, it is quite close to the optimum throughput.

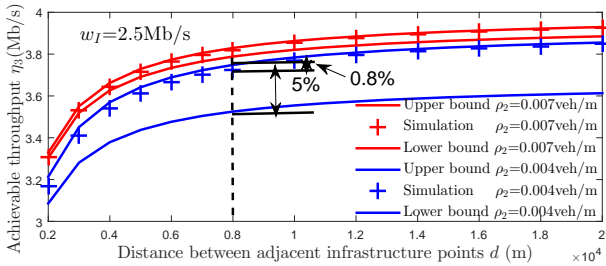
Interestingly, Fig. 5(a) and 5(c) show that in the Infrastructure-Limited regime and the transitional regime, the achievable throughput increases when  $d$  increases while Fig. 5(b) shows that in the V2V-Limited regime, the achievable throughput decreases when  $d$  increases. This can be explained that while keeping other parameters constant, an increase in  $d$  on one hand will improve the amount of data received by the VoI from V2V communications, which improves the achievable throughput; on the other hand, it will increase the amount of time spent in one *cycle* by the VoI, which reduces



(a) Infrastructure-Limited Regime



(b) V2V-Limited Regime



(c) Transitional Regime

Fig. 5. A comparison between our analytical results and the simulation results under each regime, with different *helpers'* density  $\rho_2$ .

the achievable throughput. When  $w_I$  is small, the amount of data received by the VoI from V2V communications is small due to the limitation of the amount of data received by *helpers* from infrastructure. Therefore, an increase in  $d$  will lead to a larger rate of increase in the total amount of received data by the VoI than the rate of increase in the amount of time spent in one *cycle*, which results in the overall increase of the achievable throughput, shown as Fig. 5(a) and 5(c). However, when  $w_I$  is large, the amount of data the VoI can receive from V2V communications is comparatively large, an increase in  $d$  has marginal impact on the data amount received by the VoI. Therefore, an increase in  $d$  will lead to a smaller rate of increase in the total amount of received data by the VoI than the rate of increase in the amount of time spent in one *cycle*, which results in the overall decrease of the achievable throughput, shown as Fig. 5(b).

Fig. 5 also gives insight into the optimum choice of distance between infrastructure points. It is obvious from these figures that when  $d$  increases beyond a certain threshold, e.g.,  $d=10$ km in our case, an increase in  $d$  has limited impact on the achievable throughput. This can be explained by the fact that when  $d$  is small ( $d < 10$ km in the simulation), the amount of data received by the VoI from V2V communications

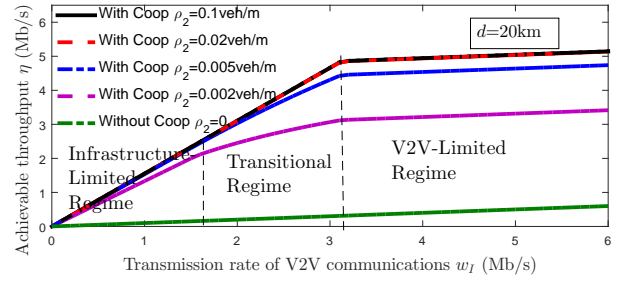


Fig. 6. A comparison between the throughput achieved from vehicular networks with and without cooperative communication by setting *helpers'* density  $\rho_2$  as 0.1veh/m (near-capacity), 0.02veh/m (congested), 0.005veh/m (low density), 0.002veh/m (low density) and 0 (without cooperation) respectively.

is relatively small compared with that received from V2I communications, especially when traffic density is low (here average  $\rho_2=0.005$ veh/m). It follows that the VoI's achievable throughput is mainly dominated by its V2I communications. However, with the increase of  $d$ , the increase of data received from V2V communications makes V2I communication's dominating impact subdued, which in turn leads to the subtle variation in the throughput.

Fig. 6 compares the achievable throughput (the lower bound for the transitional-regime is used) using our cooperative communication strategy (labeled as With coop) with its non-cooperative counterpart (labeled as Without coop). The non-cooperative counterpart is conducted by setting the *helpers'* density  $\rho_2 = 0$  because when there is no *helpers* in the vehicular network, there will be no cooperative communications. It is shown that even when *helpers'* density is low, e.g.,  $\rho_2=0.002$ veh/m, the throughput achieved by utilizing our cooperative communication strategy is around 15 times larger when  $w_I=3$ Mb/s and around 10 times larger when  $w_I=6$ Mb/s than that achieved without cooperative communications. This gives an important conclusion that our cooperative communication strategy can significantly improve the throughput even when vehicular density is rather low.

Fig. 6 also reveals the relationship between the achievable throughput and *helpers'* density  $\rho_2$ . Importantly, we can see that a higher density is beneficial to the throughput because a higher  $\rho_2$  will enhance the connectivity of vehicular networks, which leads to higher chance of V2V cooperative communications. However, when  $\rho_2$  increases beyond a certain threshold, e.g.,  $\rho_2=0.005$ veh/m in this case, a further increase in  $\rho_2$  has only marginal impact on the achievable throughput. This is due to the fact that when  $\rho_2$  is large enough for the VoI to find at least one *helper* in its coverage at any time point, increasing the density (which will lead to more *helpers* within the VoI's coverage at one time) is no longer helpful to improve the throughput because the VoI can only receive data from one vehicle at one time and the total amount of time the VoI can receive data from V2V communication will be the same.

Fig. 7 compares the achievable throughput assuming our proposed cooperative communication strategy (labeled as Coop) with that assuming the cooperative strategy proposed in [11] (labeled as ChainCluster) in the V2V-Limited regime. Specifically, the strategy proposed in [11] utilized vehicles

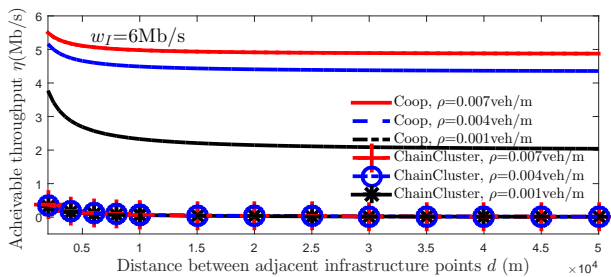


Fig. 7. A comparison between the throughput achieved from our proposed strategy and that from the strategy proposed in [11].

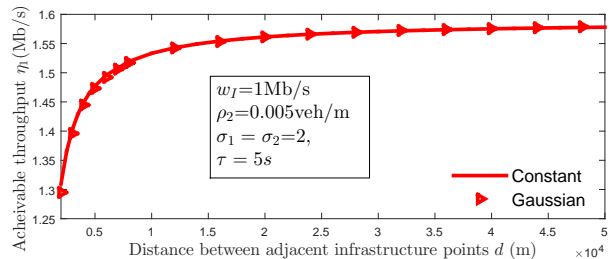


Fig. 8. A comparison between throughput achieved from the constant speed model and the time-varying speed model which follows Gaussian distribution with mean value  $v_1 = 15\text{m/s}$  and  $v_2 = 25\text{m/s}$ , variance  $\sigma_1 = \sigma_2 = 2$ , and the speed-change time interval  $\tau=5\text{s}$ .

moving in the same direction as the target vehicle (VoI) to form a cluster to help the VoI's download. A vehicle can be chosen into the cluster if and only if it can connect to the VoI via a multi-hop path. It can be seen that the throughput achieved by the VoI assuming our cooperative communication strategy is much larger than that achieved assuming the strategy proposed in [11]. This is due to the fact that in [11], the authors only used the cooperation among vehicles moving in the same direction and within the same cluster of the VoI, while in our strategy, both cooperation among infrastructure and cooperation of all vehicles traveling in the opposite direction of the VoI are fully utilized to help the VoI's download, which significantly improves the achievable throughput of the VoI.

Fig. 8 shows a comparison of the achievable throughput from the constant speed model (labeled as Constant Speed) and the time-varying speed model (labeled as Gaussian Speed) under the Infrastructure-Limited regime. The time-varying speed of vehicles in each lane follows Gaussian distributions, defined as:  $v'_1 \sim N(v_1, \sigma_1^2)$  and  $v'_2 \sim N(v_2, \sigma_2^2)$ , where  $v_1$  and  $v_2$  are the constant speed we used in our analysis, and  $\sigma_1^2$  and  $\sigma_2^2$  are the mean variance of the mean speed  $v_1$  and  $v_2$  respectively. To model the slight deviations from the mean speed, we set  $\sigma_1=\sigma_2=2$  and the speed-change time interval  $\tau=5\text{s}$ . The figure shows that when individual vehicular speed deviates slightly from the mean speed, it has marginal impact on the achievable throughput.

Fig. 9 gives a comparison of throughput achieved assuming the unit disk model (labeled as UDM) and that assuming the log-normal connection model (labeled as LNM) in the V2V-Limited regime, and shows that the unit disk model assumption has little impact on the throughput. The parameters of log-

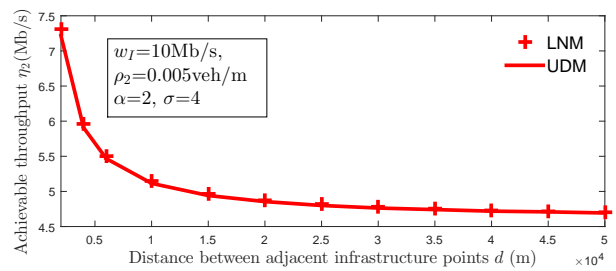


Fig. 9. A comparison between throughput achieved from the unit disk model and the log-normal connection model with path loss exponent  $\alpha=2$  and standard variance  $\sigma=4$ .

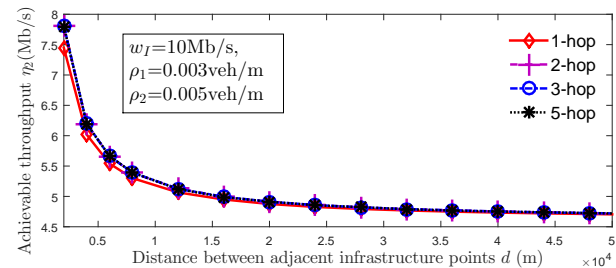


Fig. 10. A comparison between throughput achieved when allowing one-hop communication and multi-hop communications.

normal connection model are set as: path loss exponent  $\alpha=2$  and standard deviation  $\sigma=4$  [28]. It is shown that the system assuming the log-normal connection model has a slightly higher achievable throughput than that assuming the unit disk model, which coincides with the results in our previous paper [28] that log-normal connection model is beneficial to information delivery in vehicular networks. The reason behind this phenomenon is that the log-normal connection model introduces a Gaussian variation of the transmission range around the mean value, which implies a higher chance for the VoI to be connected to *helpers* separated further away.

Fig. 10 compares the throughput achieved by allowing only one-hop communication and allowing both  $k$ -hop ( $k = 2, 3, 5$ ) V2I communications between the VoI and infrastructure and  $k$ -hop V2V communication between the VoI and *helpers*. It is shown that allowing multi-hop communications beyond one hop has little impact on the throughput. Particularly, as pointed out in the end of Section III-B, in our considered scenario, allowing multi-hop V2V communication only helps to balance the distribution of information among *helpers* but do not increase the net amount of information available in the network. The marginal increase in the achievable throughput comes from multi-hop V2I communications between the VoI and infrastructure, because it allows the VoI having longer connection time with the infrastructure.

Fig. 11 compares throughput achieved from the constant channel model with that from the time-varying channel model, and shows that our analysis under the constant channel model is applicable to a more realistic time-varying channel model which considers both fading and path loss. Specifically, for the time-varying channel model, we adopt the model used in [16] that considers Rayleigh fading and path loss, from which the



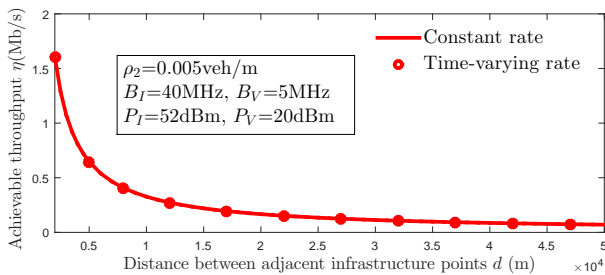


Fig. 11. A comparison between throughput achieved from constant channel model and time-varying channel model which considering Rayleigh fading and path loss.

transmission rate is given by  $w'_I = B_I \log_2 (1 + P_I |\beta d_i^{-2}|^2)$  and  $w'_V = B_V \log_2 (1 + P_V |\beta d_{ij}^{-2}|^2)$ , with the bandwidth and transmit power of each infrastructure and vehicle being  $B_I=40\text{MHz}$ ,  $P_I=52\text{dBm}$  and  $B_V=5\text{MHz}$ ,  $P_V=20\text{dBm}$  [37] respectively. Parameter  $\beta$  is the Gaussian random variable with mean 0 and variance 1 and  $d_i$ ,  $d_{ij}$  are the distances between a vehicle and its associated infrastructure point, between vehicle and vehicle when conducting V2I and V2V communications respectively. The above settings of  $B_I$ ,  $P_I$ ,  $B_V$  and  $P_V$  implies that the network is in the V2V-Limited regime. By dividing the total coverage length of the transmitter (infrastructure or vehicle) into  $K$  (here we set  $K=1000$ ) small segments, the average channel throughput  $w_I$  and  $w_V$  in the time-varying channel model can be obtained by averaging the transmission rates of all segments. This obtained average throughput  $w_I$  and  $w_V$  are then used in our constant channel model. It is obvious from Fig. 11 that the achievable throughput from the above two channel models match each other. This phenomenon can be explained by equation (35) which shows that the achievable throughput in V2V-Limited regime is a linear function of  $w_I$  and  $w_V$ . Then it follows that  $E[\eta(w_I, w_V)] = \eta(E[w_I], E[w_V])$ , which implies that for time-varying channels, the time-varying values of  $w'_I$  and  $w'_V$  can be replaced by the respective time-averaged throughput of V2I and V2V communications and our analysis still applies.

## VI. CONCLUSIONS

This paper proposed a cooperative communication strategy for vehicular networks with a finite vehicular density by utilizing V2I communications, V2V communications, mobility of vehicles, and cooperations among vehicles and infrastructure to improve the throughput. A detailed analysis for the achievable throughput was presented and the closed-form expression of achievable throughput (or its upper and lower bound) was obtained in three different regimes we classified in our analysis based on the relationship between the data rates of V2I communications, V2V communications, and the speeds of vehicles. Numerical and simulation results show that the proposed cooperative strategy can significantly improve the achievable throughput of vehicular networks even when traffic density is rather low. Simulation results show that our analysis can be extended to more realistic models such as time-varying speed model, log-normal shadowing model and time-varying channel model considering fading and path loss.

Our results shed insight on the optimum design of vehicular network infrastructure and the design of optimum cooperative communication strategies in finite vehicular networks to maximize the throughput.

## REFERENCES

- [1] Y. Li, X. Zhu, D. Jin and D. Wu, "Multiple Content Dissemination in Roadside-Unit-Aided Vehicular Opportunistic Networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 3947-3956, Oct. 2014.
- [2] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, "A Comprehensive Survey on Vehicular Ad Hoc Network," *J. Netw Comput Appl.*, vol. 37, pp. 380-392, 2014.
- [3] N. Lu, N. Cheng, N. Zhang, X. Shen and J. W. Mark, "Connected Vehicles: Solutions and Challenges," *IEEE Internet Things J.*, vol 1, no. 4, pp. 289-299, Aug. 2014.
- [4] J. Chen, A. Zafar, G. Mao, C. Li. "On the Achievable Throughput of Cooperative Vehicular Networks," in *Proceedings of IEEE ICC 2016*.
- [5] M. Khabazian, S. Aissa, and M. Mehmet-Ali, "Performance Modeling of Safety Messages Broadcast in Vehicular Ad Hoc Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 380-387, Mar. 2013.
- [6] 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.
- [7] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388-404, 2000.
- [8] M. Grossglauser and D. Tse, "Mobility increase the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477-486, Aug. 2002.
- [9] G. Mao, Z. Lin, X. Ge, and Y. Yang, "Towards a Simple Relationship to Estimate the Capacity of Static and Mobile Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 3883-3895, 2013.
- [10] M. Wang, et al., "Asymptotic Throughput Capacity Analysis of VANETs Exploiting Mobility Diversity," *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 4187-4202, Sep. 2015.
- [11] Z. Haibo, et al., "ChainCluster: Engineering a Cooperative Content Distribution Framework for Highway Vehicular Communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no.6, pp. 2644-2657, Dec. 2014.
- [12] W. Zhu, D. Li, and W. Saad, "Multiple Vehicles Collaborative Data Download Protocol via Network Coding," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1607-1619, Apr. 2015.
- [13] Y. Zhu, L. Song, S. Wu, H. Wang and C. Wang, "Cooperative Stepwise Relaying and Combining for Multihop Vehicular Wireless Communication," *IEEE Trans. Veh. Technol.*, vol 64, no. 6, pp. 2663-2671, Jun. 2015.
- [14] Q. Yan, M. Li, Z. Yang, W. Lou, and H. Zhai, "Throughput analysis of cooperative mobile content distribution in vehicular network using symbol level network coding" *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 484-492, Feb. 2012.
- [15] M. Li, Z. Yang, and W. Lou, "CodeOn: Cooperative Popular Content Distribution for Vehicular Networks using Symbol Level Network Coding," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 223-235, Jan. 2011.
- [16] T. Wang, L. Song and Z. Han, "Coalitional Graph Games for Popular Content Distribution in Cognitive Radio VANETs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 4010-4019, Oct. 2013.
- [17] D. Zhang and C. Yeo, "Enabling efficient WiFi-based vehicular content distribution," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 3, pp. 479-492, Mar. 2013.
- [18] R. Kim, H. Lim, and B. Krishnamachari, "Prefetching-Based Data Dissemination in Vehicular Cloud Systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 292 - 306, Jan. 2016.
- [19] K. Merzhad, H. Artail, and M. Gerla, "We Can Deliver Messages to Far Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1099-1115, Sep. 2012.
- [20] K. Abboud and W. Zhuang, "tochastic Analysis of a Single-Hop Communication Link in Vehicular Ad Hoc Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2297-2307, Oct. 2014.
- [21] W. Zhang, et al., "Multi-Hop Connectivity Probability in Infrastructure-Based Vehicular Networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 740-747, May. 2012.
- [22] N. Wisitpongphan, B. Fan, 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.

- [23] R. Mao and G. Mao, "Road traffic density estimation in vehicular networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013.
- [24] W. Wei, S. S. Liao, L. Xin, and J. S. Ren, "The Process of Information Propagation Along a Traffic Stream Through Intervehicle Communication," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 345-354, Feb. 2014.
- [25] Z. Zhang, G. Mao, and B. D. O. 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.
- [26] X. Ge, S. Tu, T. Han, Q. Li and G. Mao, "Energy Efficiency of Small Cell Backhaul Networks Based on Gauss-Markov Mobile Models," *IET Networks*, Vol. 4, No. 2, pp. 158-167, 2015.
- [27] L. Du, and H. Dao, "Information Dissemination Delay in Vehicle-to-Vehicle Communication Networks in a Traffic Stream", *IEEE Trans. Intell. Transp. Syst.*, vol 16, no. 1, pp. 66-80, Feb. 2015.
- [28] Z. Zhang, G. Mao, T. Han and B. D. O. Anderson, "Cooperative Information Forwarding in Vehicular Networks Subject to Channel Randomness," in *Proceedings of IEEE ICC*, 2014.
- [29] G. Mao, B. D. O. Anderson, and B. Fidan, "Online Calibration of Path Loss Exponent in Wireless Sensor Networks," in *IEEE Globecom*, 2006.
- [30] L. Tan, W. Lau, O. Yue and T. Hui, "Analytical Models and Performance Evaluation of Drive-thru Internet Systems," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 207-222, Jan. 2011.
- [31] G. Mao, and B. D. O. Anderson, "Graph Theoretic Models and Tools for the Analysis of Dynamic Wireless Multihop Networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2009.
- [32] J. Cheng, J. Cheng, M. Zhou, F. Liu, S. Gao, and C. Liu, "Routing in Internet of Vehicles: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2339-2352, Oct. 2015.
- [33] A. Ashish and P. R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," *ACM SIGCOMM Computer Communications Review*, vol. 34, no. 3, pp. 71-81, Jul. 2004.
- [34] P. Wang, G. Mao, Z. Li, X. Ge and B. D. O. Anderson, "Network Coding based Wireless Broadcast with Performance Guarantee," *IEEE Trans. Wireless Commun.*, Vol. 14, No. 1, pp. 532-544, Jan. 2015.
- [35] R. G. Gallager, *Stochastic Processes: Theory for Applications*. Cambridge University Press, 2013.
- [36] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume II*, 1971.
- [37] K. Zheng, F. Liu, Q. Zheng, W. Xiang and W. Wang, "A Graph-Based Cooperative Scheduling Scheme for Vehicular Networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1450-1458, May. 2013.

PLACE  
PHOTO  
HERE

**Guoqiang Mao** (S'98-M'02-SM'08) joined the University of Technology Sydney in February 2014 as Professor of Wireless Networking and Director of Center for Real-time Information Networks. Before that, he was with the School of Electrical and Information Engineering, the University of Sydney. The Center is among the largest university research centers in Australia in the field of wireless communications and networking. He has published about 200 papers in international conferences and journals, which have been cited more than 4500 times. He is an editor of the IEEE Transactions on Wireless Communications (since 2014), IEEE Transactions on Vehicular Technology (since 2010) and received "Top Editor" award for outstanding contributions to the IEEE Transactions on Vehicular Technology in 2011, 2014 and 2015. He is a co-chair of IEEE Intelligent Transport Systems Society Technical Committee on Communication Networks. He has served as a chair, co-chair and TPC member in a large number of international conferences. His research interest includes intelligent transport systems, applied graph theory and its applications in telecommunications, Internet of Things, wireless sensor networks, wireless localization techniques and network performance analysis.

PLACE  
PHOTO  
HERE

**Changle Li** (M'09) received the Ph.D. degrees in communication and information system from Xidian University, China, in 2005. Since then, he conducted his postdoctoral research in Canada and the National Institute of Information and Communications Technology (NICT), Japan, respectively. He has been a visiting scholar at the University of Technology Sydney (UTS) and is currently a Professor with the State Key Laboratory of Integrated Services Networks, Xidian University. He is an IEEE Senior Member and his research interests include intelligent transportation systems, vehicular networks, mobile ad hoc networks, and wireless sensor networks.

PLACE  
PHOTO  
HERE

**Ammar Zafar** obtained his PhD in Wireless Communications from King Abdullah University of Science and Technology under the supervision of Dr. Mohamed-Slim Alouini in 2014. In November 2014, he joined Centre of Real-time Information Networks at the University of Technology Sydney as a post-doctoral research fellow. Ammar Zafar's research interests include traffic flow monitoring, vehicular networks, multi-user scheduling and wireless communication theory.

PLACE  
PHOTO  
HERE

**Albert Y. Zomaya** is the Chair Professor of High Performance Computing & Networking in the School of Information Technologies, University of Sydney, and he also serves as the Director of the Centre for Distributed and High Performance Computing. Professor Zomaya published more than 600 scientific papers and articles and is author, co-author or editor of more than 20 books. He is the Founding Editor in Chief of the IEEE Transactions on Sustainable Computing and serves as an associate editor for more than 20 leading journals. Professor Zomaya served as an Editor in Chief for the IEEE Transactions on Computers (2011-2014). Professor Zomaya is the recipient of the IEEE Technical Committee on Parallel Processing Outstanding Service Award (2011), the IEEE Technical Committee on Scalable Computing Medal for Excellence in Scalable Computing (2011), and the IEEE Computer Society Technical Achievement Award (2014). He is a Chartered Engineer, a Fellow of AAAS, IEEE, and IET. Professor Zomaya's research interests are in the areas of parallel and distributed computing and complex systems.

PLACE  
PHOTO  
HERE

**Jieqiong Chen** (S'16) received the Bachelor's degree in Engineering from Zhejiang University, Zhejiang, China, in 2012, and she is currently working toward the Ph.D. degree in engineering at the University of Technology Sydney, Sydney, Australia. Her research interests in the area of wireless communications and intelligent transportation systems.