

# Cooperative Content Transmission for Vehicular Ad Hoc Networks using Robust Optimization

Daxin Tian\*, Jianshan Zhou\*, Min Chen†, Zhengguo Sheng‡, Qiang Ni§ and Victor C.M. Leung¶

\*School of Transportation Science and Engineering, Beihang University, Beijing 100191, China

Email: dtian@buaa.edu.cn, jianshanzhou@foxmail.com

†School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China

Email: minchen2012@hust.edu.cn

‡Department of Engineering and Design, University of Sussex, Richmond 3A09, UK

Email: z.sheng@sussex.ac.uk

§School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, UK

Email: q.ni@lancaster.ac.uk

¶Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, B.C., V6T 1Z4 Canada

Email: vleung@ece.ubc.ca

**Abstract**—Vehicular ad hoc networks (VANETs) have a potential to promote vehicular telematics and infotainment applications, where a key and challenging issue is the design of robust and efficient vehicular content transmissions to combat the lossy inter-vehicle links. In this paper, we focus on the robust optimization of content transmissions over cooperative VANETs. We first derive a stochastic model for estimation of time-varying inter-vehicle distance, which is dependent of the vehicle real-time kinematics and the distribution of the initial space headway. With this model, we analytically formulate the transient inter-vehicle connectivity assuming Nakagami fading channels for the physical (PHY) layer. We also model the contention nature of the medium access control (MAC) layer, on which we are based to evaluate the throughput achieved by each vehicle equipped with dedicated short-range communication (DSRC). Combining these models, we derive a closed-form expression for the upper bound of the probability of failure in intact-content transmissions. Based upon this theoretical bound, we develop a robust optimization model for assigning content data traffic among different cooperative transmission paths, where the objective is to minimize the maximum likelihood of unsuccessful content transmissions over the cooperative VANET. We mathematically transform the optimization model to another equivalent form, such that it can be practically deployed. Finally, we validate our theoretical development with extensive simulations. Numerical results are also provided to confirm the power of cooperation in boosting the VANET performance as well as demonstrate the advantage of the proposed robust optimization in terms of content data reception reliability.

## I. INTRODUCTION

VEHICULAR ad-hoc networks (VANETs) have drawn considerable research interest, because they are envisioned to support a diverse array of vehicular applications with different quality of service (QoS) requirements. Different from short message disseminations in VANETs [1], vehicular content transmissions aim to achieve bulk data transfer between fast moving vehicles equipped with DSRC. The content files may contain different types of information, such as text messages, images, video and audio data, which can be divided into a series of data packets in heterogeneous smaller sizes. At this point, VANETs-based content transmissions representatively

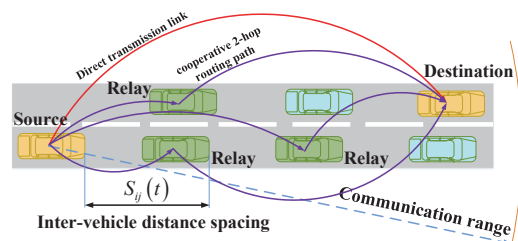


Fig. 1. A typical scenario of cooperative vehicular content transmissions.

boil down to the process of consecutively routing data packets of different sizes from a source's upper-layer applications to a destination's [2]–[4]. To guarantee the QoS or the QoE of vehicle users, their upper-layer multimedia applications always require the integrity of content data received from their transmitters [2], [3], [5], [6]. The consecutive deliveries of data packets may rely on multiple cooperative routing paths [7]–[10]. However, it is non-trivial to achieve intact-content transmissions in the inter-vehicle channel. Several common significant characteristics of cooperative VANETs will make this design issue a challenging task [11], such as *Dynamic and stochastic topology* [12], *Fast V2V channel fading* [13], [14], and *Serious channel contention* [15].

Towards this end, in this work, our focus is on the design of a reliable and efficient content transmission scheme for cooperative VANETs, using the robust optimization philosophy [16]. Our design philosophy exploits the stochastic analysis from a cross-layer perspective and the maximin mechanism as a powerful robust modeling tool for the treatment of uncertainty and randomness in cooperative VANETs-based content transmissions. Specifically, a system model for inter-vehicle throughput estimation is first developed by combining contention nature in the MAC layer with the stochastic models of the time-varying inter-vehicle spacing, the channel fading and the PHY-layer modulation rate adaptation. Then, we present a stochastic analysis on the throughput performance and derive a

closed-form expression for bounding the possibility of failure in content transfer through an inter-vehicle link. Finally, we develop a robust optimization model equipped with the adaptive relay selection, so as to determine the optimal assignment of heterogeneous-size data packet-level traffic over the feasible cooperative routing paths as well as the direct transmission path between the source and its destination. This model aims to minimize the upper bound of the likelihood of unsuccessful content transmissions over the whole vehicular routing paths, meanwhile achieving the optimal payload balance among these paths. Using mathematical analysis, we derive another equivalent formation of the optimization model that is more suitable for implementation. Finally, extensive simulations are conducted to validate our proposed stochastic models and prove the effectiveness of our optimization method. To the best of our knowledge, the proposed framework presents the first effort in literature to facilitate the robust optimization of cooperative vehicular content transmissions with consideration of complex coupled effects arising from vehicle mobility, channel contentions and fading.

The rest of this paper is organized as follows. In Section II, we develop different stochastic models and then arrive at the model for the throughput estimation of cooperative VANETs. Section III proposes the robust optimization model for realization of optimal cooperative content transmissions. Section IV evaluates our theoretical development. Finally, Section V concludes this paper.

## II. SYSTEM MODEL

### A. Modeling of time-varying inter-vehicle distance

We consider a common traffic scenario where the vehicles equipped with wireless communication, i.e., DSRC, and the unequipped are mixed while they are moving with the homogeneous kinematic model. We begin our modeling with assuming that after the connection setup, i.e., after the time instant  $t_0 = 0$ , the time-varying inter-vehicle distance between any two selected transmitter and receiver, vehicles  $i$  and  $j$ , is a random variable  $S_{i,j}(t)$  as shown in Fig. 1. Let the speeds of  $i$  and  $j$  at the initial time  $t_0$  be  $v_i$  and  $v_j$ , respectively, and their instantaneous accelerations be  $A_i$  and  $A_j$ , respectively. Because of stochastic fluctuations in the traffic flow and noises in the vehicle dynamics systems such as braking system and power train system, these accelerations may not be constant. Considering this reality, we express  $A_i = a_i + \eta_i$  and  $A_j = a_j + \eta_j$  where  $a_i$  and  $a_j$  are two mean accelerations during the short-lived connection between  $i$  and  $j$ , while  $\eta_i$  and  $\eta_j$  are assumed two independently identically distributed random noises. Since the duration of vehicular transmissions is usually much shorter than that of the state transition in traffic flows, the vehicle's motion can be approximated by

$$\begin{cases} S_i(t) = v_i t + \frac{1}{2} A_i t^2, \\ S_j(t) = v_j t + \frac{1}{2} A_j t^2, \end{cases} \quad (1)$$

where  $S_i(t)$  and  $S_j(t)$  are the distances traveled by  $i$  and  $j$  by the time  $t$ , respectively. It should be noticed that  $v_i, v_j, a_i$  and

$a_j$  are directional parameters. We assume that the direction of vehicle  $j$  is positive, i.e.,  $v_j \geq 0$ , so that when  $i$  and  $j$  are moving in the same direction, we see  $v_i \geq 0$ ; otherwise, when  $i$  is moving in the opposite direction,  $v_i < 0$ . The accelerations  $a_i$  and  $a_j$  are defined positive if and only if they are in the same direction of  $v_j$ .

Denoting  $v_{i,j} = v_j - v_i$  and  $a_{i,j} = a_j - a_i$ , we obtain

$$\begin{aligned} S_{i,j}(t) &= (S_j(t) - S_i(t)) + s_{i,j} \\ &= v_{i,j}t + \frac{1}{2}a_{i,j}t^2 + s_{i,j} + \frac{1}{2}(\eta_j - \eta_i)t^2 \end{aligned} \quad (2)$$

in which  $s_{i,j}$  is the initial inter-vehicle spacing at  $t_0$  between vehicles  $i$  and  $j$ . In addition, we assume that the noises are characterized by a Gaussian distribution with the mean value  $\mu_\eta$  and the standard  $\sigma_\eta$ , i.e.,  $\eta_i, \eta_j \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$ . Let their PDF be  $g_\eta(x)$ . By algebraic transformation, we can further get  $g_{-\eta}(y) = g_\eta(-y)$ .

We introduce another random variable  $Z = \eta_j - \eta_i$ , the realization of which is  $z = x + y$ . Recalling that  $\eta_i$  and  $\eta_j$  are independent with each other and  $g_{-\eta}(y) = g_\eta(-y)$ , we yield the two-dimensional joint PDF of  $Z$  by  $g_Z(x, y) = g_\eta(x)g_{-\eta}(y) = g_\eta(x)g_\eta(-y)$ , which results in

$$g_Z(z) = \int_{-\infty}^{+\infty} g_Z(x, z-x)dx = \int_{-\infty}^{+\infty} g_\eta(x)g_\eta(x-z)dx. \quad (3)$$

Then, under the assumption of the Gaussian distribution for the PDF  $g_\eta(x)$ , the detailed expression of (3) can be

$$g_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_\eta^2} \exp\left\{-\frac{(x-\mu_\eta)^2 + (x-z-\mu_\eta)^2}{2\sigma_\eta^2}\right\} dx. \quad (4)$$

For simplicity, we represent the deterministic term in (2) by  $s_{i,j}(t) = v_{i,j}t + \frac{1}{2}a_{i,j}t^2 + s_{i,j}$ , such that the space headway  $S_{i,j}(t)$  can be re-expressed as  $S_{i,j}(t) = s_{i,j}(t) + \frac{1}{2}t^2Z$ . Using the result of (4), we further derive the PDF of  $S_{i,j}(t)$  over time  $t$  conditioned upon a collection of  $i$ 's and  $j$ 's kinematic parameters,  $\theta_{i,j} = \{v_i, v_j, a_i, a_j\}$ , and the initial inter-vehicle spacing,  $s_{i,j}$ , as follows

$$g_{S_{i,j}(t)}(d|\theta_{i,j}, s_{i,j}) = \frac{2}{t^2} g_Z\left(\frac{2}{t^2}(d - s_{i,j}(t))\right). \quad (5)$$

From (5), the PDF of the time-varying  $S_{i,j}(t)$  indeed incorporates the impacts of the vehicle mobility and the initial inter-vehicle spacing distribution.

### B. Modeling of vehicular channel fading

We assume that the communication vehicles in a mixed traffic flow are equipped with DSRC/IEEE802.11p, which propagate the PHY-layer signal on the licensed spectrum of 75MHz in the frequency band from 5.850 to 5.925 GHz [17]–[19]. According to much existing literature [6], [13], [15], [19], we adopt the Nakagami distribution to capture the fast small-scale fading in vehicular propagation channels. Let  $P$  be the power received at the receiver  $j$  from the transmitter  $i$ , and the corresponding received signal envelope from  $i$  be  $\sqrt{P}$ , which

is a random variable following a Nakagami distribution with the parameters  $(m, \omega)$ . Thus, the PDF of  $\sqrt{P}$  is

$$g_{\sqrt{P}}(x | m, \omega) = \frac{2m^m}{\Gamma(m)\omega^m} x^{2m-1} \exp\left(-\frac{m}{\omega} x^2\right) \quad (6)$$

where  $\Gamma(m)$  is a Gamma distribution function, i.e.,  $\Gamma(m) = \int_0^{+\infty} x^{m-1} e^{-x} dx$ .  $m$  is the channel fading parameter, and  $\omega$  is the expected received power in the fading envelope. Based on equation (6), the CDF of the received power  $P$  is

$$\text{Prob}(P \leq x) = \frac{\gamma\left(m, \frac{m}{\omega} x\right)}{\Gamma(m)} \quad (7)$$

where the Gamma function  $\gamma\left(m, \frac{m}{\omega} x\right)$  is represented by  $\gamma\left(m, \frac{m}{\omega} x\right) = \int_0^{\frac{m}{\omega} x} s^{m-1} e^{-s} ds$ .

To calculate  $\omega$ , we refer to the measurements reported in [19]. To be specific, we employ a dual-slope piecewise-linear model to approximate the signal strength,  $P(d)$ , received by  $j$  from  $i$  given the inter-vehicle distance  $d = |S_{i,j}(t)|$ :

$$P(d) = \begin{cases} P(d_0) - 10\gamma_1 \log_{10}\left(\frac{d}{d_0}\right) + X_{\sigma_1}, & d_0 \leq d \leq d_c; \\ P(d_0) - 10\gamma_1 \log_{10}\left(\frac{d_c}{d_0}\right) \\ \quad - 10\gamma_2 \log_{10}\left(\frac{d}{d_c}\right) + X_{\sigma_2} & , d > d_c; \end{cases} \quad (8)$$

where  $\gamma_1$  and  $\gamma_2$  are two path loss exponents, while  $X_{\sigma_1}$  and  $X_{\sigma_2}$  are two independently random variables following zero-mean normal distributions with the standard deviations  $\sigma_1$  and  $\sigma_2$ , respectively.  $P(d_0)$  is a given signal power received at a reference distance  $d_0$ .  $d_c$  is referred to as the Fresnel distance. It can be approximated by the formula  $d_c = \frac{4H_i H_j}{\lambda_{\text{wave}}}$ , where  $H_i$  and  $H_j$  are the antenna heights of vehicles  $i$  and  $j$ , respectively, and  $\lambda_{\text{wave}}$  is the wavelength of the electromagnetic wave at 5.9GHz [19]. Thus, the expected received power in the fading envelope given the inter-vehicle distance  $S_{i,j}(t)$  is  $\omega = \omega(S_{i,j}(t)) = \mathbb{E}[P(d) | d = |S_{i,j}(t)|]$ .

Furthermore, we exploit a distance-based model to evaluate the channel fading parameter  $m$ . According to the real-world measurements in [19], the fading in vehicular propagation channel can be characterized with a finite set of different parameters,  $\{a_k, k = 1, 2, \dots, K\}$ . Each fading parameter  $a_k$  corresponds to a certain range of the transmission distance,  $\mathcal{D}_k = (S_{k-1}, S_k]$ . That is, when the inter-vehicle distance,  $d$ , is ranging within  $\mathcal{D}_k$ , i.e.,  $d \in \mathcal{D}_k$ ,  $m$  can be set to the corresponding  $a_k$ , i.e.,  $m = a_k$  (here we set the lower bound of the first distance interval,  $\mathcal{D}_1$ , by  $S_0 = 0$ ). We notice that the condition  $d = |S_{i,j}(t)| \in \mathcal{D}_k$  is equivalent to the union of  $-S_k \leq S_{i,j}(t) \leq -S_{k-1}$  and  $S_{k-1} \leq S_{i,j}(t) \leq S_k$ . Thus, with the stochastic model (5), we can derive the probability mass function of the channel fading parameter  $m$  conditioned on  $d = |S_{i,j}(t)|$  as follows

$$\text{Prob}(m = a_k) = \int_{-S_k}^{-S_{k-1}} g_{S_{i,j}(t)}(x | \theta_{i,j}, s_{i,j}) dx + \int_{S_{k-1}}^{S_k} g_{S_{i,j}(t)}(x | \theta_{i,j}, s_{i,j}) dx \quad (9)$$

for  $k = 1, 2, \dots, K$ .

Now, given the power of the thermal noise at  $j$ ,  $I$ , the SNR is defined by  $R = \frac{P}{I}$ . Using (7), we obtain the CDF of the SNR level at  $j$  given the channel fading parameter  $m$  and the average signal strength  $\omega(S_{i,j}(t))$

$$\text{Prob}(R \leq x | m, \omega(S_{i,j}(t))) = \frac{\gamma\left(m, \frac{m}{\omega(S_{i,j}(t))} Ix\right)}{\Gamma(m)}. \quad (10)$$

Following the result (10), the connectivity of a transmission pair, defined as the probability  $p_{\text{con}}$  that the SNR at the receiver  $j$  is larger than a specified threshold  $\phi$  to correctly receive information, can be easily determined by

$$p_{\text{con}} = \text{Prob}(R > \phi | m, \omega(S_{i,j}(t))) = 1 - \frac{\gamma\left(m, \frac{m}{\omega(S_{i,j}(t))} I\phi\right)}{\Gamma(m)}. \quad (11)$$

### C. Modeling of vehicular channel contentions

Let  $\lambda_g$  represent the overall vehicle density in vehicles per meter and  $\lambda_e$  be the market penetration rate,  $0 < \lambda_e \leq 1$ . In addition, we assume that only a part of vehicles among the equipped vehicles are willing to serve as cooperative relays, and let  $p_c$  denote the cooperation ratio,  $0 < p_c \leq 1$ . A random variable  $N$  denotes the number of the cooperative equipped vehicles that would contend to access the same channel for communications. Following much existing literature such as [1], [6], we assume that the presence of the cooperative equipped vehicles in a mixed traffic flow follows an independent homogeneous spatial Poisson process with intensity  $\lambda = \lambda_g \lambda_e p_c$ . Hence, the probability mass function of  $N$  can be presented by

$$\text{Prob}(N = n) = \frac{(S\lambda)^n}{n!} \exp(-S\lambda), \quad (12)$$

where  $S$  is the carrier sensing range of any equipped vehicle.

We assume that the cooperative VANET employs a contention-based access mechanism in the MAC layer to resolve the vehicular channel contentions. Specifically, we consider that the RTS/CTS (Request-to-Send and Clear-to-Send) scheme is used for eliminating the hidden terminals as well as the IEEE 802.11 distributed coordination function (DCF) for MAC-layer scheduling. Denote by  $W$  the contention window size of the backoff process in the MAC layer and  $\tau$  the average transmission probability that an equipped vehicle transmits in an idle slot. Applying the mean approximation immediately obtains

$$\tau = \frac{2}{W + 2}. \quad (13)$$

Therefore, with (13), the probability of a communication vehicle encountering an idle slot can be expressed as

$$p_{\text{idl}} = (1 - \tau)^N, \quad (14)$$

and that of success in a transmission is

$$p_{\text{suc}} = \binom{N}{N-1} \tau (1 - \tau)^{N-1}. \quad (15)$$

Besides, the probability of a collided transmission occurring can be determined by  $p_{\text{col}} = 1 - p_{\text{idl}} - p_{\text{suc}}$ .

#### D. Throughput performance of per-transmission link

To evaluate the MAC-layer throughput performance of the transmission link between the transmitter  $i$  and the receiver  $j$ , we denote by  $L_i$  the average data payload of vehicle  $i$ ,  $L_{\text{head}}$  the packet header including the PHY-layer header and the MAC-layer header, i.e.,  $L_{\text{head}} = L_{\text{PHY}} + L_{\text{MAC}}$ . The unit time slot in the MAC-layer DCF scheme is denoted by  $\delta_{\text{slot}}$ . Additionally, let  $\delta_{\text{DIFS}}$  (the time interval of DCF Interframe Space),  $\delta_{\text{SIFS}}$  (the time interval of Short Interframe Space),  $\delta_{\text{RTS}}$ ,  $\delta_{\text{CTS}}$  and  $\delta_{\text{ACK}}$  be the pre-specified time intervals reserved for the DCF-related operations (such as signalings) and the transmissions of RTS-, CTS- and ACK- messages. According to the IEEE 802.11 standard and with these notations, we can express the average time of collided transmission,  $\delta_{\text{col}}$ , and that of successful transmission,  $\delta_{\text{suc}}$ , as follows

$$\begin{cases} \delta_{\text{col}} = \delta_{\text{RTS}} + \delta_{\text{DIFS}} + \delta_{\text{slot}}; \\ \delta_{\text{suc}} = \delta_{\text{RTS}} + \delta_{\text{DIFS}} + \delta_{\text{CTS}} + \delta_{\text{ACK}} \\ \quad + 3\delta_{\text{SIFS}} + 4\delta_{\text{slot}} + \frac{\mathbb{E}[L]}{\mathbb{E}[C]}, \end{cases} \quad (16)$$

where  $\mathbb{E}[L]$  denotes the average packet length of  $N$  vehicles contending the wireless channel and  $\mathbb{E}[C]$  the average transmission rate among these vehicles.

To evaluate the average transmission rate of the cooperative VANET, we resort to the discrete SNR-based channel modulation scheme, in which the vehicular communication terminal is assumed to support a finite set of  $Q$  different PHY-layer modulation rates,  $\{c_r, r = 1, 2, \dots, Q\}$ . These modulation rates are adopted according to the SNR level. That is, each modulation rate,  $c_r$ , is triggered when the SNR level,  $R$ , is ranging within a certain range  $(\phi_r, \phi_{r+1}]$ , i.e.,  $R \in (\phi_r, \phi_{r+1}]$ . Based on equations (10) and (11), we can derive the probability of the selection of the modulation rate  $c_r$  by

$$\begin{aligned} \text{Prob}(C = c_r | m, \omega(S_{i,j}(t))) = \\ \begin{cases} \frac{\gamma(m, \frac{m}{\omega(S_{i,j}(t))} I\phi_{r+1}) - \gamma(m, \frac{m}{\omega(S_{i,j}(t))} I\phi_r)}{\Gamma(m)}, r = 1, \dots, Q-1; \\ 1 - \frac{\gamma(m, \frac{m}{\omega(S_{i,j}(t))} I\phi_Q)}{\Gamma(m)}, r = Q; \end{cases} \end{aligned} \quad (17)$$

Hence, we can get the theoretical MAC-layer throughput,  $u_{i,j}$ ,

$$u_{i,j} = \frac{\tau p_{\text{suc}}(L_i + L_{\text{head}})}{p_{\text{idl}}\delta_{\text{slot}} + p_{\text{col}}\delta_{\text{col}} + p_{\text{suc}}\delta_{\text{suc}}}. \quad (18)$$

Following equation (18), given a limited time interval  $[0, T]$ , the total data volume that can be completely transmitted from vehicle  $i$  to  $j$  within the specified time duration  $T$  can be expressed as follows

$$U_{i,j} = \int_0^T u_{i,j} dt. \quad (19)$$

### III. ROBUST OPTIMIZATION OVER COOPERATIVE VANETS

#### A. Analysis of expected per-vehicle throughput

From (18), it can be seen that the throughput  $U_{i,j}$  is a random variable. This fact renders it impractical to compute  $U_{i,j}$  from (19). Thus, we would like to evaluate the expected throughput performance, i.e.,

$$\mathbb{E}[U_{i,j}] = \int_0^T \mathbb{E}[u_{i,j}] dt. \quad (20)$$

Next, we derive the main theoretical results.

**Result 1:** Given the kinematics of the transmitter  $i$  and the receiver  $j$ ,  $\theta_{i,j}$ , as well as the initial inter-vehicle spacing,  $s_{i,j}$ , the expected MAC-layer throughput performance of the transmission pair from  $i$  to  $j$  is as follows

$$\mathbb{E}[U_{i,j}] = \sum_{n=1}^{+\infty} \sum_{k=1}^K \sum_{r=1}^Q \int_0^T \int_{-\infty}^{+\infty} u_{i,j} p(c_r, a_k, n, x) dx dt, \quad (21)$$

where  $p(c_r, a_k, n, x)$  is defined by

$$\begin{aligned} p(c_r, a_k, n, x) = & \text{Prob}(C = c_r, m = a_k, N = n, S_{i,j}(t) = x) \\ = & \text{Prob}(C = c_r | m, \omega(x)) \\ & \times \text{Prob}(m = a_k) \times \text{Prob}(N = n) \\ & \times g_{S_{i,j}}(x | \theta_{i,j}, s_{i,j}). \end{aligned} \quad (22)$$

*Proof:* **Result 1** is based on equations (5), (9), (12), and (17), and naturally follows the mathematical definition of the conditional expectation and the integral linearity. ■

Denote by  $f_{i,j}$  the size of the content file, i.e., the data volume, expected to be forwarded from  $i$  to  $j$  within the given time duration  $[0, T]$ . We define the *negative reliability* associated with the transmission pair  $(i, j)$  as  $\text{Prob}(U_{i,j} \leq f_{i,j})$ , which indicates the possibility of the assigned content file size exceeding the per-vehicle throughput performance, i.e., the likelihood of failing in content transmissions. Thus, we provide the following theorem to approximate its upper bound.

**Theorem 1:** Given the kinematics of the transmitter  $i$  and the receiver  $j$ ,  $\theta_{i,j}$ , the initial inter-vehicle spacing,  $s_{i,j}$ , as well as the size of the content file to be transmitted,  $f_{i,j}$ , and  $f_{i,j} \leq \mathbb{E}[U_{i,j}]$ , the probability of a vehicular transmission link,  $(i, j)$ , failing to achieve the content transmission within a given duration  $[0, T]$  has an upper bound  $\text{Prob}_{\text{upper}}(f_{i,j} | \theta_{i,j}, s_{i,j})$ :

$$\begin{aligned} \text{Prob}_{\text{upper}}(f_{i,j} | \theta_{i,j}, s_{i,j}) = \\ \frac{T \int_0^T \mathbb{E}[u_{i,j}^2] dt - (\mathbb{E}[U_{i,j}])^2}{T \int_0^T \mathbb{E}[u_{i,j}^2] dt - (\mathbb{E}[U_{i,j}])^2 + (\mathbb{E}[U_{i,j}] - f_{i,j})^2}. \end{aligned} \quad (23)$$

*Proof:* The proof can follow the same way in [6]. ■

#### B. Robust optimization

We follow much current literature to assume that each vehicle equipped with DSRC also deploys a GPS onboard and is running a routine service to exchange some control messages

among its neighbors. Thus, the other neighbor can access the kinematics and the real-time location of the host vehicle, and can evaluate the inter-vehicle distance as well. Certain computing information used for decision making can also be locally exchanged through the routine service. Denote the set of cooperative relays between  $i$  and  $j$  by  $\mathcal{V}_{i,j}$ , and the set consisting of all the cooperative two-hop routing paths and the direct path by  $\mathcal{P}_{i,j}$ , i.e.,  $\mathcal{P}_{i,j} = \{[(i, j)], [(i, v_p), (v_p, j)], v_p \in \mathcal{V}_{i,j}\}$ .  $i$  divides the overall content data  $F_{i,j}$  into  $|\mathcal{P}_{i,j}|$  data traffic streams with each being routed over one path in  $\mathcal{P}_{i,j}$ . Moreover, we let a fragment of the content file  $F_{i,j}$ , i.e., the data traffic assigned to a routing path  $p \in \mathcal{P}_{i,j}$ , be  $f_p \geq 0$  and the collection of all  $f_p$  be  $\mathbf{f} = \text{col}\{f_p, \forall p \in \mathcal{P}_{i,j}\}$ . Thus, we can define a cooperation cost function in terms of end-to-end transmission reliability as

$$G(\mathbf{f}) = \sum_{p \in \mathcal{P}_{i,j}} f_p h_p(f_p), \quad (24)$$

where the path-related cost function  $h_p(f_p)$  is formulated as follows, which quantifies the likelihood of the cooperative path  $p$  failing in transmission of the assigned content data  $f_p$ :

$$h_p(f_p) = 1 - \prod_{e \in p} (1 - \text{Prob}(U_e \leq f_p)). \quad (25)$$

Here, we use  $e \in p$  to represent a transmission link associated with the cooperative routing path  $p$ . For example, given a  $p = [(i, v_p), (v_p, j)] \in \mathcal{P}_{i,j}$ , the intermediate relay  $v_p$  connects to two links,  $e_1, e_2 \in p$ , one of which is  $e_1 = (i, v_p)$  and the other  $e_2 = (v_p, j)$ . Intuitively, the optimization of dividing the large-data-volume content file into a series of heterogeneous-size manageable chunks and assigning them to multiple cooperative routing paths can be modeled as follows

$$\begin{aligned} \min_{\mathbf{f}} : G(\mathbf{f}) &= \sum_{p \in \mathcal{P}_{i,j}} f_p h_p(f_p) \\ \text{s.t.} \quad &\begin{cases} F_{i,j} = \sum_{p \in \mathcal{P}_{i,j}} f_p; \\ f_p \leq \frac{1}{\beta} \min\{\mathbb{E}[U_e], e \in p\}, p \in \mathcal{P}_{i,j}; \\ 0 \leq f_p, p \in \mathcal{P}_{i,j}, \end{cases} \end{aligned} \quad (26)$$

where  $\beta \in (0, 1]$  is a tunable parameter. For the sake of simplicity, we denote the set of constraints in (26) by  $\mathcal{F}_{i,j}$ .

Indeed, (26) is traditionally classified as a probabilistic optimization model. However, due to the unknown distribution of  $U_e$ , i.e., existing uncertainty in each component,  $\text{Prob}(U_e \leq f_p)$ , of the cooperation cost function  $G(\mathbf{f})$  given above, it is impractical to solve this probabilistic model. In this situation, it is meaningful to make a robust decision from the conservative point of view. Hence, we resort to the robust optimization paradigm, i.e., the dominating paradigm in this area of robust optimization, to propose a robust counterpart. Namely, we further formulate the following *minimax* model

$$\begin{aligned} \min_{\mathbf{f}} \max_{\mathbf{f}} : G(\mathbf{f}) &= \sum_{p \in \mathcal{P}_{i,j}} f_p h_p(f_p) \\ \text{s.t.} \quad &\mathbf{f} \in \mathcal{F}_{i,j}. \end{aligned} \quad (27)$$

Based on the result of **theorem 1**, we can easily see

$$h_p(f_p) \leq 1 - \prod_{e \in p} [1 - \text{Prob}_{\text{upper}}(f_p | \boldsymbol{\theta}_e, s_e)]. \quad (28)$$

Let the right term of the inequality (28) above be  $H_{\text{upper}}(f_p | \boldsymbol{\theta}_e, s_e)$ . Accordingly, with (28), we transform the minmax optimization model (27) into

$$\begin{aligned} \min_{\mathbf{f}} : W(\mathbf{f}) &= \sum_{p \in \mathcal{P}_{i,j}} f_p H_{\text{upper}}(f_p | \boldsymbol{\theta}_e, s_e) \\ \text{s.t.} \quad &\mathbf{f} \in \mathcal{F}_{i,j}. \end{aligned} \quad (29)$$

From (29), it is seen that this robust optimization model involves both the equality and the inequality constraints. Thus, it is non-trivial to directly solve (29). To cope with (29), we first consider to solve a subproblem similar to (29) while only dealing with the equality constraint, i.e.,

$$\begin{aligned} \min_{\mathbf{f}} : W_1(\mathbf{f}) &= \sum_{p \in \mathcal{P}_{i,j}} f_p H_{\text{upper}}(f_p | \boldsymbol{\theta}_e, s_e) \\ \text{s.t.} \quad &F_{i,j} - \mathbf{1}^T \mathbf{f} = 0, \end{aligned} \quad (30)$$

where we define  $\mathbf{1}$  as a column vector with size  $|\mathcal{P}_{i,j}| \times 1$ , whose entries are all identical to 1, i.e.,  $\mathbf{1} = \text{col}\{1, \dots, 1\} \in \mathbb{R}^{|\mathcal{P}_{i,j}| \times 1}$ . Now, we present the following theorem based on the analysis of the submodel (30).

**Theorem 2:** There exists a nonnegative scalar parameter  $\alpha'_1 \geq 0$ , such that obtaining a strict local optimal solution of (30) boils down to solve a strict local minimum point of the following unconstrained optimization problem:

$$\begin{aligned} \min_{\mathbf{f}, \mu_1} : L_1(\mathbf{f}, \mu_1) &= W_1(\mathbf{f}) - \mu_1 (F_{i,j} - \mathbf{1}^T \mathbf{f}) \\ &\quad + \frac{\alpha_1}{2} (F_{i,j} - \mathbf{1}^T \mathbf{f})^2, \end{aligned} \quad (31)$$

under the condition that the factor  $\alpha_1$  satisfies  $\alpha_1 > \alpha'_1$ .

*Proof:* Let the solution  $\bar{\mathbf{f}}$  satisfy the properties of a strict local optimal solution of the subproblem (30). That is, there exists a multiplier  $\bar{\mu}_1$  that makes the following conditions held

$$\begin{cases} \nabla_{\mathbf{f}} (W_1(\bar{\mathbf{f}}) - \bar{\mu}_1 (F_{i,j} - \mathbf{1}^T \bar{\mathbf{f}})) = \nabla_{\mathbf{f}} W_1(\bar{\mathbf{f}}) + \bar{\mu}_1 \mathbf{1} = 0; \\ F_{i,j} - \mathbf{1}^T \bar{\mathbf{f}} = 0. \end{cases} \quad (32)$$

Besides, according to the strict optimality of  $\bar{\mathbf{f}}$ , it satisfies the second-order sufficient condition:

$$\mathbf{y}^T \nabla_{\mathbf{f}}^2 (W_1(\bar{\mathbf{f}}) - \bar{\mu}_1 (F_{i,j} - \mathbf{1}^T \bar{\mathbf{f}})) \mathbf{y} = \mathbf{y}^T \nabla_{\mathbf{f}}^2 W_1(\bar{\mathbf{f}}) \mathbf{y} > 0 \quad (33)$$

is always held for any nonzero column vector  $\mathbf{y} \in \mathbb{R}^{|\mathcal{P}_{i,j}| \times 1}$  that satisfies  $\mathbf{y}^T \nabla_{\mathbf{f}} (F_{i,j} - \mathbf{1}^T \bar{\mathbf{f}}) = -\mathbf{y}^T \mathbf{1} = 0$ .

Now, we take the partial derivative of the objective function in (31) with respect to  $\mathbf{f}$  so as to get

$$\nabla_{\mathbf{f}} L_1(\mathbf{f}, \mu_1) = \nabla_{\mathbf{f}} W_1(\mathbf{f}) + \mu_1 \mathbf{1} - \alpha_1 (F_{i,j} - \mathbf{1}^T \mathbf{f}) \mathbf{1}. \quad (34)$$

Substituting the conditions of (32) into (34) above can obtain  $\nabla_{\mathbf{f}} L_1(\bar{\mathbf{f}}, \bar{\mu}_1) = 0$ . Thus,  $\bar{\mathbf{f}}$  and  $\bar{\mu}_1$  also meet the first-order condition of the unconstrained optimization model (31).

Next, based on (34), we further yield the Hessian matrix  $\nabla_{\mathbf{f}}^2 L_1(\mathbf{f}, \mu_1)$  as follows

$$\nabla_{\mathbf{f}}^2 L_1(\mathbf{f}, \mu_1) = \nabla_{\mathbf{f}}^2 W_1(\mathbf{f}) + \alpha_1 \mathbf{1}\mathbf{1}^T. \quad (35)$$

Recalling  $\alpha_1 > \alpha'_1 \geq 0$ , we can easily see that

$$\begin{aligned} \mathbf{y}^T \nabla_{\mathbf{f}}^2 L_1(\bar{\mathbf{f}}, \bar{\mu}_1) \mathbf{y} &= \mathbf{y}^T (\nabla_{\mathbf{f}}^2 W_1(\bar{\mathbf{f}}) + \alpha_1 \mathbf{1}\mathbf{1}^T) \mathbf{y} \\ &> \mathbf{y}^T \nabla_{\mathbf{f}}^2 W_1(\bar{\mathbf{f}}) \mathbf{y} > 0 \end{aligned} \quad (36)$$

is held due to (33). At this point, the Hessian matrix  $\nabla_{\mathbf{f}}^2 L_1(\bar{\mathbf{f}}, \bar{\mu}_1)$  at the point  $\bar{\mathbf{f}}$  is strictly positively definite when  $\alpha_1 > 0$ . Combining the results of  $\nabla_{\mathbf{f}} L_1(\bar{\mathbf{f}}, \bar{\mu}_1) = 0$  and the strict positive definiteness, we arrive at the conclusion that  $\bar{\mathbf{f}}$  is also a strict local minimum point of the unconstrained optimization model (31).

Conversely, suppose that a point  $\tilde{\mathbf{f}}$  meets  $F_{i,j} - \mathbf{1}^T \tilde{\mathbf{f}} = 0$  as well as is a strict local minimum point of (31) with a certain parameter  $\tilde{\mu}_1$ , i.e.,  $\nabla_{\mathbf{f}} L_1(\tilde{\mathbf{f}}, \tilde{\mu}_1) = 0$ . These conditions directly lead to

$$\nabla_{\mathbf{f}} W_1(\tilde{\mathbf{f}}, \tilde{\mu}_1) + \tilde{\mu}_1 \nabla_{\mathbf{f}} (F_{i,j} - \mathbf{1}^T \tilde{\mathbf{f}}) = 0, \quad (37)$$

which indicates that  $\tilde{\mathbf{f}}$  is a Karush-Kuhn-Tucker (KKT) point of the model (30).

Similar to (35), we can also derive  $\nabla_{\mathbf{f}}^2 L_1(\tilde{\mathbf{f}}, \tilde{\mu}_1) = \nabla_{\mathbf{f}}^2 W_1(\tilde{\mathbf{f}}) + \alpha_1 \mathbf{1}\mathbf{1}^T$ . Due to the strict optimality of  $\tilde{\mathbf{f}}$ , we have  $\tilde{\mathbf{y}}^T \nabla_{\mathbf{f}}^2 L_1(\tilde{\mathbf{f}}, \tilde{\mu}_1) \tilde{\mathbf{y}} > 0$  for any nonzero column vector  $\tilde{\mathbf{y}}$ , which is equivalent to

$$\tilde{\mathbf{y}}^T (\nabla_{\mathbf{f}}^2 W_1(\tilde{\mathbf{f}}) + \alpha_1 \mathbf{1}\mathbf{1}^T) \tilde{\mathbf{y}} > 0. \quad (38)$$

Therefore, (37) and (38) indicate that  $\tilde{\mathbf{f}}$  is also a strict local optimal solution of the model (30).

In summary, the derivations above show that we can solve a strict local minimum point of the subproblem (31) to obtain a strict local optimal solution of the model (30). ■

Another fact revealed in **Theorem 2** is that we are allowed to solve the subproblem (31) for a strict local minima with any fixed finite parameter  $\alpha_1 > 0$ . Thus, we can pre-specify  $\alpha_1$  and do not need to ensure  $\alpha_1 \rightarrow +\infty$ . This makes great sense in designing an optimization algorithm for (31).

Now, we turn to deal with the original problem (29) under both the equality and the inequality constraints. Specifically, we present the following result based on **Theorem 2**.

**Theorem 3:** There exists a nonnegative scalar parameter  $\alpha'_2 \geq 0$ , such that obtaining a strict local optimal solution of (29) boils down to solve a strict local minimum point of the unconstrained optimization problem:

$$\begin{aligned} \min_{\mathbf{f}, \mu_2, \boldsymbol{\pi}, \boldsymbol{\kappa}} : L_2(\mathbf{f}, \mu_2, \boldsymbol{\pi}, \boldsymbol{\kappa}) &= W_1(\mathbf{f}) - \mu_2 (F_{i,j} - \mathbf{1}^T \mathbf{f}) \\ &+ \frac{\alpha_2}{2} (F_{i,j} - \mathbf{1}^T \mathbf{f})^2 \\ &+ \frac{1}{2\alpha_2} \sum_{p \in \mathcal{P}_{i,j}} \left( (\xi_1(\pi_p, f_p))^2 - \pi_p^2 \right) \\ &+ \frac{1}{2\alpha_2} \sum_{p \in \mathcal{P}_{i,j}} \sum_{e \in p} \left( (\xi_2(\kappa_{p,e}, f_p))^2 - \kappa_{p,e}^2 \right), \end{aligned} \quad (39)$$

under the condition  $\alpha_2 > \alpha'_2$ , where  $\boldsymbol{\pi} = \{\pi_p, p \in \mathcal{P}_{i,j}\}$  and  $\boldsymbol{\kappa} = \{\kappa_{p,e}, e \in p, p \in \mathcal{P}_{i,j}\}$  are two different sets of Lagrange multipliers.  $\xi_1(\pi_p, f_p)$  and  $\xi_2(\kappa_{p,e}, f_p)$  are given by

$$\begin{cases} \xi_1(\pi_p, f_p) = \max\{\alpha_2 f_p - \pi_p, 0\} \\ \xi_2(\kappa_{p,e}, f_p) = \max\left\{\alpha_2 \left(\frac{1}{\beta} \mathbb{E}[U_e] - f_p\right) - \kappa_{p,e}, 0\right\}. \end{cases} \quad (40)$$

*Proof:* To prove this theorem, we introduce two sets of slack variables, i.e.,  $\mathbf{w}_1 = \{w_p, p \in \mathcal{P}_{i,j}\}$  and  $\mathbf{w}_2 = \{w_{p,e}, e \in p, p \in \mathcal{P}_{i,j}\}$ , so as to transform the inequality constraints in  $\mathcal{F}_{i,j}$  into the equality constraints:  $f_p \geq 0 \rightarrow f_p - w_p^2 = 0$  for  $\forall p \in \mathcal{P}_{i,j}$ , and  $f_p \leq \frac{1}{\beta} \max\{\mathbb{E}[U_e], e \in p\} \rightarrow \frac{1}{\beta} \mathbb{E}[U_e] - f_p - w_{p,e}^2 = 0$  for  $\forall e \in p$  and  $\forall p \in \mathcal{P}_{i,j}$ . Thus, the model (29) can be rearranged as

$$\begin{aligned} \min_{\mathbf{f}, \mathbf{w}_1, \mathbf{w}_2} : W_1(\mathbf{f}) &= \sum_{p \in \mathcal{P}_{i,j}} f_p H_{\text{upper}}(f_p | \boldsymbol{\theta}_e, s_e) \\ \text{s.t.} \begin{cases} F_{i,j} - \mathbf{1}^T \mathbf{f} = 0; \\ f_p - w_p^2 = 0, p \in \mathcal{P}_{i,j}; \\ \frac{1}{\beta} \mathbb{E}[U_e] - f_p - w_{p,e}^2 = 0, e \in p, p \in \mathcal{P}_{i,j}. \end{cases} \end{aligned} \quad (41)$$

Based on **Theorem 2**, solving (41) is equivalent to solve

$$\begin{aligned} \min_{\mathbf{f}} : \tilde{L}_2(\mathbf{f}, \mathbf{w}_1, \mathbf{w}_2, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa}) &= W_1(\mathbf{f}) \\ &- \mu_2 (F_{i,j} - \mathbf{1}^T \mathbf{f}) + \frac{\alpha_2}{2} (F_{i,j} - \mathbf{1}^T \mathbf{f})^2 \\ &- \sum_{p \in \mathcal{P}_{i,j}} \pi_p (f_p - w_p^2) + \frac{\alpha_2}{2} \sum_{p \in \mathcal{P}_{i,j}} (f_p - w_p^2)^2 \\ &- \sum_{p \in \mathcal{P}_{i,j}} \sum_{e \in p} \kappa_{p,e} \left( \frac{1}{\beta} \mathbb{E}[U_e] - f_p - w_{p,e}^2 \right) \\ &+ \frac{\alpha_2}{2} \sum_{p \in \mathcal{P}_{i,j}} \sum_{e \in p} \left( \frac{1}{\beta} \mathbb{E}[U_e] - f_p - w_{p,e}^2 \right)^2. \end{aligned} \quad (42)$$

As for (42), we treat it as an unconstrained optimization problem with respect to  $\{\mathbf{w}_1, \mathbf{w}_2\}$ , such that we can solve it for the minimum points  $\{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2\}$ . To achieve this, we mathematically rearrange the cost function in (42) as follows

$$\begin{aligned} \tilde{L}_2(\mathbf{f}, \mathbf{w}_1, \mathbf{w}_2, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa}) &= W_1(\mathbf{f}) \\ &- \mu_2 (F_{i,j} - \mathbf{1}^T \mathbf{f}) + \frac{\alpha_2}{2} (F_{i,j} - \mathbf{1}^T \mathbf{f})^2 \\ &+ \sum_{p \in \mathcal{P}_{i,j}} \left\{ \frac{\alpha_2}{2} \left[ w_p^2 - \frac{1}{\alpha_2} (\alpha_2 f_p - \pi_p) \right]^2 - \frac{\pi_p^2}{2\alpha_2} \right\} \\ &+ \sum_{p \in \mathcal{P}_{i,j}} \left\{ \frac{\alpha_2}{2} \left[ w_{p,e}^2 - \frac{1}{\alpha_2} \left( \alpha_2 \left( \frac{1}{\beta} \mathbb{E}[U_e] - f_p \right) - \kappa_{p,e} \right) \right]^2 \right\} \\ &- \sum_{p \in \mathcal{P}_{i,j}} \sum_{e \in p} \frac{\kappa_{p,e}^2}{2\alpha_2}. \end{aligned} \quad (43)$$

It is obvious that if and only if  $w_p$  and  $w_{p,e}$  are set to

$$\begin{cases} \tilde{w}_p^2 = \frac{1}{\alpha_2} \xi_1(\pi_p, f_p), \\ \tilde{w}_{p,e}^2 = \frac{1}{\alpha_2} \xi_2(\kappa_{p,e}, f_p), \end{cases} \quad (44)$$

for  $\forall e \in p$  and  $\forall p \in \mathcal{P}_{i,j}$ , the cost function (43) attains its local minimum value, i.e.,  $L_2(\mathbf{f}, \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa})$ . Therefore, we can set  $L_2(\mathbf{f}, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa}) = \tilde{L}_2(\mathbf{f}, \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa})$ , such that the local optimal solution of the model (29) with respect to  $\mathbf{f}$  can be obtained by further minimizing  $L_2(\mathbf{f}, \mu_2, \alpha_2, \boldsymbol{\pi}, \boldsymbol{\kappa})$ . At this point, we complete the proof. ■

We remark that **Theorem 3** indeed provides an equivalent model for the original robust model (29), which is much easier to be tackled and more practical for computing. Currently, there have been many numerical unconstrained optimization algorithms, such as the steepest descent method, the conjugate gradient method, and the quasi-newton algorithm, etc., on which we can rely to solve the equivalent robust model (40) for an optimal data traffic assignment policy for cooperative vehicular content transmissions.

#### IV. SIMULATIONS AND NUMERICAL RESULTS

##### A. Simulation scenario and parameters

We consider a linear highway scenario where vehicles equipped with DSRC and the other unequipped coexist. Vehicles are randomly mixed in a traffic flow, forming a spatial Poisson point process. To validate our theoretical models proposed before, we develop our simulations using Python-based object-oriented programming technique. In our simulations, some basic PHY-layer and MAC-layer simulation parameters of each vehicle with DSRC are given according to some current studies [6], [15], [19], as provided in Table I. It should be noted that the carrier sensing range is set to much larger than that of the one-hop range so as to suppress the hidden terminal issue. The time resolution of the mobility simulation is set to  $dt = 0.1s$ , while that of the communication simulation should be  $1\mu s$  as specified in IEEE 802.11 standard. Furthermore, we also adopt different distance-based channel fading parameters  $m$  for the Nakagami channels and different received SNR-based PHY-layer transmission rates  $C$ . We remark that the simulation settings adopted here is for the sake of demonstration of how to apply our mathematical models. Nevertheless, our models are independent of a specific parameter in Table I. Once the PHY-layer or MAC-layer parameters in other scenarios of interest are appropriately defined, our model can be implemented as well.

##### B. Distribution of stochastic inter-vehicle spacing

To validate the stochastic model of inter-vehicle spacing in (5), we vary the general traffic density  $\lambda_g$ , i.e.,  $\lambda_g \in [1/200, 1/100, 1/50]$ veh/m and fix the market penetration rate  $\lambda_e = 0.8$  and the cooperation ratio  $p_c = 0.9$ . Thus, the initial inter-vehicle distance can be set to  $1/(\lambda_g \lambda_e p_c)$ m/veh. In the simulated traffic flow, we select two equipped vehicles, vehicles  $i$  and  $j$ , with initial  $v_i = 90/3.6$ m/s,  $v_j = 95/3.6$ m/s,  $a_i = 1.0$ m/s<sup>2</sup>, and  $v_j = -1.0$ m/s<sup>2</sup>. The mean and the standard deviation of the distribution of the stochastic acceleration noise  $\eta$  are  $\sigma_\eta = 0$ m/s<sup>2</sup> and  $0.2(m/s^2)^2$ , respectively. We perform Monte Carlo simulations of the mobility scenario and then compare the simulation results with the results of our theoretical model. All the simulations have been performed

TABLE I  
BASIC PHY-LAYER AND MAC-LAYER PARAMETERS.

Parameters in the propagation model (8)	Values
Path loss exponents $\gamma_1, \gamma_2$	2.1, 3.8
Standard deviations $\sigma_1, \sigma_2$	2.6dB, 4.4dB
Critical distance $d_c$	100m
Reference distance $d_0$	38.31m
Reference power $P(d_0)$	-38.8325dBm
Parameters in the SNR CDF model (10)	Values
Thermal noise power $I$	-96dBm
Parameters in channel contentions	Values
Contention window size $W$	32
Header of PHY-layer frame $L_{\text{PHY}}$	136Bytes
Header of MAC-layer frame $L_{\text{MAC}}$	28Bytes
Unit time slot $\delta_{\text{slot}}$ and ACK slot $\delta_{\text{ACK}}$	13 $\mu s$ , 37 $\mu s$
RTS and CTS intervals $\delta_{\text{RTS}}, \delta_{\text{CTS}}$	53 $\mu s$ , 37 $\mu s$
DIFS and SIFS intervals $\delta_{\text{DIFS}}, \delta_{\text{SIFS}}$	58 $\mu s$ , 32 $\mu s$
PDF of the packet payload distribution	$\mathcal{N}(500, 83.3^2)$ Bytes
One-hop communication range	300.0m
Carrier sensing range $S$	500.0m
Distance interval	$m$
$d \in (0, 100]$ m	$a_1 = 1.0$
$d \in (100, 200]$ m	$a_2 = 0.5$
$d \in (200, 500]$ m	$a_3 = 0.1$
SNR interval	$C$
$R \in (0, 25]$ dB	$c_1 = 12.0$ Mbps
$R \in (25, +\infty)$ dB	$c_2 = 24.0$ Mbps

with 200 replications per simulation point (i.e., per  $s_{i,j}$ ), such that the results corresponding to each  $s_{i,j}$  are illustrated with a statistical histogram.

Fig. 2 compares the results of computing the model (5) and the simulation results when the travel time  $t$  is set to  $t = 20s$  and the initial inter-vehicle spacing  $s_{i,j}$  is set to different values. As can be seen, the PDF curve obtained by the model (5) can excellently approximate the statistics of the simulation results. Additionally, Fig. 2 shows that a larger initial inter-vehicle spacing is more likely to result in a larger inter-vehicle distance at the time  $t$  in the stochastic mobility scenario.

##### C. Analysis of connectivity probability

In this subsection, we further investigate the connectivity of the cooperative VANET under the Nakagami Fading Channel. To achieve this, we set the SNR threshold in the model (11) by  $\phi = 3$ dB. The traffic density  $\lambda_g$  and the market penetration rate  $\lambda_e$  are as the same as those in Subsection IV-B, while  $p_c$  is discretely increased from 0.1 to 1.0 with a step 0.0474.

Fig. 3 shows this experiment outcomes where the markers denote the mean of simulation results and the curves are theoretical results. As is demonstrated, we can increase  $p_c$  or  $\lambda_g$  to improve the connectivity probability of a transmission pair. The comparison between the theoretical results and those of the simulations can confirm the validity of our theoretical model. To give an insight into the coupled effects of  $s_{i,j}$  and  $t$ , we evaluate the mean connectivity denoted by  $\mathbb{E}[p_{\text{con}}]$ , i.e.,  $\mathbb{E}[p_{\text{con}}] = \int_{-\infty}^{+\infty} p_{\text{con}} g_{S_{i,j}}(t)(x|\boldsymbol{\theta}_{i,j}, s_{i,j})dx$ . The numerical results are illustrated in Fig. 4. From this figure, we can see that a smaller inter-vehicle spacing and a shorter travel time trend to increase the mean connectivity probability. Indeed, there exists an optimal region of  $s_{i,j}$  and  $t$  where  $\mathbb{E}[p_{\text{con}}]$



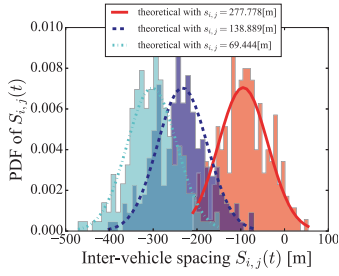


Fig. 2. PDF of  $S_{i,j}(t)$  at  $t = 20s$  under different densities  $\lambda$ .

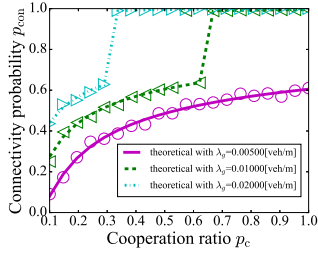


Fig. 3. Connectivity probability  $p_{con}$  under different cooperation ratios  $p_c$ .

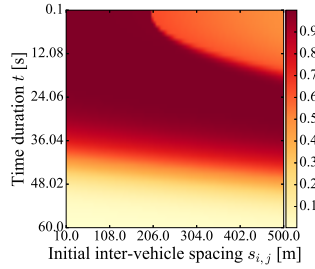


Fig. 4. Expected connectivity under different  $t$  and different  $s_{i,j}$ .

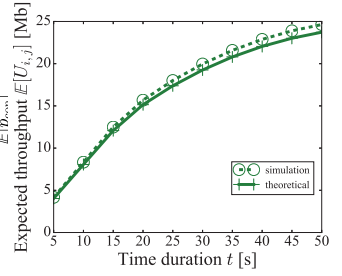


Fig. 5. Theoretical  $\mathbb{E}[U_{i,j}]$  and simulation results under different  $t$ .

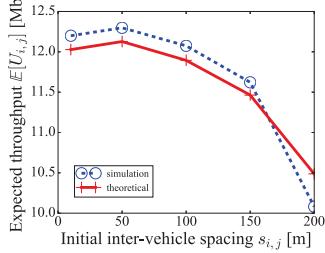


Fig. 6. Theoretical  $\mathbb{E}[U_{i,j}]$  and simulation results under different  $s_{i,j}$ .

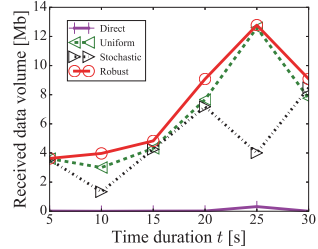


Fig. 7. Received data volume with different schemes under different  $t$ .

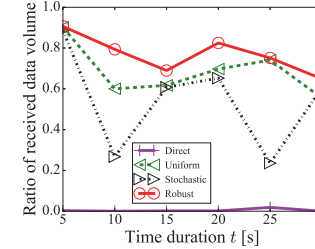


Fig. 8. Ratio of received data volume over that needed to be transmitted.

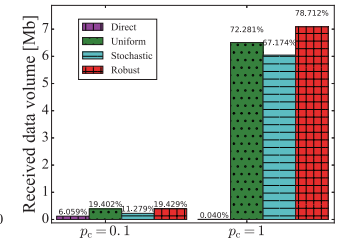


Fig. 9. Received data volume with different schemes under different  $p_c$ .

can be maintained at a high level. For example, when  $t$  does not exceed about 30.0s and  $s_{i,j}$  is smaller than about 200.0m,  $\mathbb{E}[p_{con}]$  always holds  $\mathbb{E}[p_{con}] \geq 0.8$ . Another interesting fact shown in Fig. 4 is that when  $s_{i,j}$  is large, e.g.,  $s_{i,j} > 300m$ , a smaller  $t$ , e.g.,  $t < 12s$ , cannot guarantee a good connectivity. As can be seen from Fig. 4, at a large  $s_{i,j} > 300m$ ,  $\mathbb{E}[p_{con}]$  is changed from a lower region to a higher region, and again falls into a much lower region as  $t$  increases from 0.1s to 60.0s. The main reason is that  $S_{i,j}(t)$  is directional. Namely, as  $t$  increases,  $|S_{i,j}(t)|$  decreases to 0 and then becomes larger than 0, such that the connectivity experiences both an increasing and a decreasing trends.

#### D. Transmission pair throughput

Next, we conduct simulations to validate the theoretical evaluation of the expected throughput performance (21). In the experiment, we set  $\lambda_g = 1/50veh/m$  while  $p_c = 0.9$  and  $\lambda_e = 0.8$ . We also select two transmission pair  $i$  and  $j$  for observation, whose speed, acceleration and acceleration noise are set in the same way in Subsection IV-B. The data payload of a packet is assumed to be 1000Bytes.

We vary  $t$  from 5s to 50s with an increment step of 5s. At each point of  $t$ , the simulations have been carried out with 200 replications, and the results are shown with their mean value. Fig. 5 compares the results obtained by computing (21) with those of the simulations. As can be observed, (21) provides a good approximation to the simulation results. Furthermore, we can also see that increasing  $t$  can increase the mean transmitted data volume, while the increment rate is reduced. This is because the connectivity probability of the pair would reduce when  $t$  increases, as illustrated in Fig. 4. In Fig. 6, with  $t = 15s$ , we investigate the variation of the mean data volume

transmitted under different  $s_{i,j}$ . The theoretical results can also appropriately approximate those of the simulations.

#### E. Performance of robust optimization

Finally, we compare the proposed method (denoted by “Robust”) with some other schemes, including the direct transmission scheme (“Direct”) that performs the content transmissions without cooperation, the uniform assignment scheme (“Uniform”) that uniformly assigns the content data traffic of the same volume to each cooperative path, the stochastic assignment scheme (“Stochastic”) that randomly generates an assignment solution. It is noted that except for “Direct”, the others schemes exploit the cooperation mechanism.

We randomly select a transmission pair  $i$  and  $j$  from a vehicle flow with  $\lambda_g = 1/50veh/m$ . We set  $\lambda_e = 0.9$ ,  $p_c = 0.8$ ,  $\beta = 0.1$  and  $\sigma_\eta = 0.1m/s^2$ . The speeds of all the vehicles are uniformly randomly generated from  $[90/3.6, 95/3.6]m/s$ , and their accelerations from  $[-1, 1]m/s^2$ . The time duration  $t$  is discretely varied from 5s to 30s, and the size of the total content file,  $F_{i,j}$ , is generated from  $[4, 14]Mb$ . Fig. 7 and Fig. 8 illustrate the performance with different mechanisms, showing that our method can achieve the largest volume of content data received by the destination as well as the best reception ratio. In addition, Fig. 9 shows the results under two different cooperation conditions,  $p_c = 0.1$  and  $p_c = 1$ . In Fig. 9, the ratio of data volume received by the destination is expressed as the percentage attached above the corresponding bar chart. In these two situations, we fix  $t = 15s$  and  $F_{i,j} = 2Mb, 9Mb$ , respectively. As can be found, using the cooperation can boost the received data volume by about 10% on average under  $p_c = 0.1$ , while the received data volume is increased by about 72% on average under  $p_c = 1$ . To examine



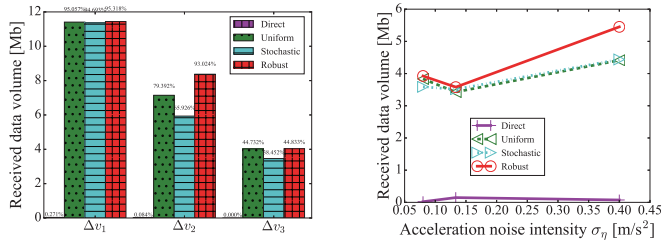


Fig. 10. Received data volume with Fig. 11. Received data volume with different schemes under different  $\Delta v$ . different schemes under different  $\sigma_\eta$ .

how the mobility impacts on the performance, we further set the speed interval, denoted by  $\Delta v$ , within which vehicle speed is randomly uniformly distributed, to different cases,  $\Delta v_1 = [30/3.6, 35/3.6]$ m/s,  $\Delta v_2 = [60/3.6, 65/3.6]$ m/s, and  $\Delta v_3 = [90/3.6, 95/3.6]$ m/s. In each case, let  $t = 15$ s and  $p_c = 0.8$ . The experimental results are compared in Fig. 10, indicating that higher mobility would reduce the received data volume. Besides, we show the effects of different  $\sigma_\eta$  on the performance in Fig. 11. We fix the uniform speed interval as  $[90/3.6, 95/3.6]$ m/s,  $F_{i,j} = 6$ Mb, and  $\sigma_\eta$  as  $\sigma_\eta \in [4/50, 4/30, 4/10]$ m/s<sup>2</sup>. From Fig. 7 to Fig. 11 above, it can be confirmed that the VANET can gain respectable benefit from the cooperation among transmitters, and the proposed method can always perform the best among the schemes with and without cooperation at all the simulation points.

## V. CONCLUSION

In this paper, we use the advantage of cooperation among interacting vehicles to form a cooperative VANET for vehicular content file transmissions. To guarantee the performance of the cooperative VANET, we develop a robust optimization framework for effective and reliable realization of cooperative vehicular content transmissions. Comprehensive effects of the vehicle mobility, contentions and fading in the inter-vehicle propagation channel are considered in our theoretical development. The proposed robust optimization method is theoretically analyzed and proved to benefit the cooperative VANET through extensive comparative simulations. Notably, this work bridges the divide between the complex coupled impacts, which arise from vehicle mobility and vehicular communications, and the robust optimization of content transmission performance in cooperative VANETs. In the future, we would like to investigate the performance of the proposed robust optimization in real-world situations by conducting field tests. Moreover, it is also expected to incorporate other more complex driving behaviors, e.g., car-following dynamics, in the inter-vehicle spacing model, and to extend the robust optimization model to the car-following scenarios.

## ACKNOWLEDGMENT

This research is supported in part by the National Natural Science Foundation of China under Grant No. 61672082 and 61711530247, the Engineering and Physical Sciences

Research Council (EPSRC) (EP/P025862/1), Royal Society-Newton Mobility Grant (IE160920) and Asa Briggs Visiting Fellowship from University of Sussex.

## REFERENCES

- [1] D. Tian, J. Zhou, Y. Wang, G. Zhang, and H. Xia, "An adaptive vehicular epidemic routing method based on attractor selection model," *Ad Hoc Networks*, vol. 36, pp. 465 – 481, 2016, vehicular Networking for Mobile Crowd Sensing. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1570870515001237>
- [2] Z. Su, Y. Hui, and Q. Yang, "The next generation vehicular networks: A content-centric framework," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 60–66, February 2017.
- [3] J. A. Khan and Y. Ghamri-Doudane, "Saving: socially aware vehicular information-centric networking," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 100–107, August 2016.
- [4] M. Gerla and L. Kleinrock, "Vehicular networks and the future of the mobile internet," *Computer Networks*, vol. 55, no. 2, pp. 457 – 469, 2011, wireless for the Future Internet. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1389128610003324>
- [5] M. A. Salkuyeh and B. Abolhassani, "An adaptive multipath geographic routing for video transmission in urban vanets," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2822–2831, Oct 2016.
- [6] T. H. Luan, X. S. Shen, and F. Bai, "Integrity-oriented content transmission in highway vehicular ad hoc networks," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 2562–2570.
- [7] Z. Ding and K. K. Leung, "Cross-layer routing using cooperative transmission in vehicular ad-hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 571–581, 2011.
- [8] D. Tian, J. Zhou, Z. Sheng, M. Chen, Q. Ni, and V. C. M. Leung, "Self-organized relay selection for cooperative transmission in vehicular ad-hoc networks," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2017.
- [9] M. Asefi, J. W. Mark, and X. Shen, "A cross-layer path selection scheme for video streaming over vehicular ad-hoc networks," in *2010 IEEE 72nd Vehicular Technology Conference - Fall*, Sept 2010, pp. 1–5.
- [10] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 18–29, Sept 2006.
- [11] H. Hartenstein and L. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164–171, June 2008.
- [12] G. Yan and S. Olariu, "A probabilistic analysis of link duration in vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1227–1236, Dec 2011.
- [13] R. He, A. F. Molisch, F. Tufvesson, Z. Zhong, B. Ai, and T. Zhang, "Vehicle-to-vehicle propagation models with large vehicle obstructions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2237–2248, Oct 2014.
- [14] M. Boban, T. T. V. Vinhoza, M. Ferreira, J. Barros, and O. K. Tonguz, "Impact of vehicles as obstacles in vehicular ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 1, pp. 15–28, January 2011.
- [15] R. Chen, Z. Sheng, Z. Zhong, M. Ni, V. C. M. Leung, D. G. Michelson, and M. Hu, "Connectivity analysis for cooperative vehicular ad hoc networks under nakagami fading channel," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1787–1790, Oct 2014.
- [16] S. Verdu and H. Poor, "On minimax robustness: A general approach and applications," *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 328–340, Mar 1984.
- [17] D. Jiang and L. Delgrossi, "Ieee 802.11p: Towards an international standard for wireless access in vehicular environments," in *VTC Spring 2008 - IEEE Vehicular Technology Conference*, May 2008, pp. 2036–2040.
- [18] S. Biswas, R. Tatchikou, and F. Dion, "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 74–82, Jan 2006.
- [19] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 ghz dedicated short range communication (dsrc) frequency band," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1501–1516, Oct 2007.