Semantic Communication-Based Dynamic Resource Allocation in D2D Vehicular Networks

Jiawei Su, Zhixin Liu, *Senior Member, IEEE*, Yuan-ai Xie, Kai Ma, Hongyang Du, Jiawen Kang, Dusit Niyato, *Fellow, IEEE*

***Abstract*—The semantic communication mechanism enables wireless devices in vehicular networks to communicate more effectively with the semantic meaning. However, in high-dynamic vehicular networks, the transmission of semantic information faces challenges in terms of reliability and stability. To address these challenges, a long-term robust resource allocation scheme is proposed under the Device-to-Device (D2D) vehicular (D2D-V) networks, where multiple performance indicators (user satisfaction, queue stability, and communication delay) are considered. Due to the sophisticated probabilistic form with consideration of channel fluctuations, the Bernstein approximation is introduced to acquire the deterministic constraint more efficiently. The robust resource allocation problem is proposed and separated into two independent subproblems by the Lyapunov optimization method, which includes semantic access control in the application layer and power control in the physical layer. After that, the successive convex approximation method and Karush-Kuhn-Tucher conditions are adopted to solve the subproblems, thereby proposing a robust resource allocation algorithm. The simulations reveal the trade-off relationship between user satisfaction, queue stability, and communication delay, which is on the premise of meeting the user SINR requirement. Moreover, the simulations also prove the necessity of considering channel uncertainty in high-speed mobile vehicular communication scenarios.**

***Index Terms*—Vehicular networks, Semantic communication, Resource allocation, Lyapunov optimization, D2D technology.**

I. Introduction

To improve traffic efficiency and driving experience for vehicle users, autonomous driving technology has been widely studied [1]. In autonomous driving systems, efficient decisions of vehicles mainly depend on the data collected from their surroundings (mainly by taking traffic pictures by cameras) and the information interaction with adjacent vehicles/ roadside infrastructures (wireless communications). Compared with the perception of surroundings, information interaction based on the internet of vehicle (IoV) is particularly critical [2]. However, efficient and reliable vehicular communications face two major technical challenges [3].

On the one hand, due to the large amount of data collected by vehicles, the transmission of original data will require large bandwidth and cause severe communication overhead, which is not realistic for current IoV with scarce spectrum resources. A novel communication mode, semantic communication mechanism, has been widely studied [4]. With the improvement of intelligence and informatization, vehicles have the ability to terminal information calculation. The improvement of terminal ability provides support for semantic communication which extracts the core semantic data and only transmits the meanings [5]. On the other hand, to facilitate the direct information interaction between adjacent vehicles, device-to-device (D2D) technology is widely used in vehicular networks and facilitates a D2D vehicular (D2D-V) system [6]. However, due to the frequent topology changes caused by the high mobility of vehicles, the vehicle-to-everything link is easy to be interrupted. Therefore, the guarantee of Signal to Interference plus Noise Ratio (SINR) is the key point in a dynamic environment [7]. Recent studies demonstrated that the semantic communication mechanism is more robust in a low SINR environment [8], which increases the communication qualities of wireless devices. Based on the extracted semantic data, D2D communications aim to maximize the transmission efficiency and minimize the semantic errors by transmitting the meaning of data, rather than traditional bit errors.

Nevertheless, since the semantic data cannot be transmitted by D2D users instantaneously, it has to be temporarily cached in the buffer of the application layer. Therefore, cross-layer resource allocation which joints the application layer and the physical layer is the most widely used method to control queue stability and improve system robustness, especially in dynamic communication environments [9]. As a result, a cross-layer robust resource allocation framework based on the D2D-V networks is a promising solution for efficient and reliable vehicular communications.

II. Related Works

According to Shannon's theorem, there is a limit value of transmission rate in theory, which represents the maximum communication capacity of the channel. As Shannon argued, converting a continuous source data requires a channel with infinite capacity, and the solution is to discretize the signal within a certain tolerance of information loss [10]. In other words, traditional communication model based on Shannon's theory has limited efficiency. Semantic communication, as a novel communication mode, has been widely studied in recent years, which extracts the semantic data and only transmits the core meanings. Instead of discretizing the continuous source

Jiawei Su, Zhixin Liu, Yuan-ai Xie and Kai Ma are with the School of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China. (e-mail: Sjw@stumail.ysu.edu.cn, lzxauto@ysu.edu.cn, xieyuan_ai@163.com, kma@ysu.edu.cn)

H. Du is with the School of Computer Science and Engineering, the Energy Research Institute at NTU, Interdisciplinary Graduate Program, Nanyang Technological University, Singapore (e-mail: hongyang001@e.ntu.edu.sg).

J. Kang is with the School of Automation, Guangdong University of Technology, China. (e-mail: kavinkang@gdut.edu.cn)

D. Niyato is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore (e-mail: dniyato@ntu.edu.sg)

signal with a certain loss tolerance, semantic communication allows transmitting the meaning of the signal, which has the potential to significantly improve transmission efficiency (lower channel resource and lossless signal delivery) [11]. Furthermore, the semantic error can be greatly reduced by transmitting the core meaning of data, which is under the help of a semantic knowledge library to encode and decode respectively [12]. However, the existing semantic communication research does not well consider the process of end-to-end data exchange, and the ultra-reliable and low-latency semantic transmission that meets various user qualities of service (QoS) has not been paid enough attention [13].

Therefore, this paper considers a joint scenario of semantic communications and D2D technology and studies the communication quality of signal links on the basis of extracted semantic data. The combination of semantic communications and D2D technology is full of advantages, i.e., reducing communication load, realizing end-to-end direct transmission, reducing transmission errors, and so on. To improve the network stability, the traditional solutions mainly focus on the power control in the physical layer [14]. However, the semantic data cannot be transmitted by D2D users instantaneously, it has to be temporarily cached in the buffer of the application layer, ignoring the access rate of semantic data in the application layer will lead to an imbalance data queue, thereby producing uncontrollable network delays. Therefore, a long-term dynamic cross-layer resource allocation framework is highlighted and constructed to guarantee the requirements of QoS and the queue stability in [9]. Lyapunov optimization method is adopted in this paper, which firstly transforms the long-term constraints into queue stability conditions and then transforms the long-term objective functions and the queue stability conditions into solvable short-term subproblems [15].

Although the Lyapunov optimization method has greater advantages in long-term performance indexes, it cannot well depress the problems of co-channel interference. As is stated in [16], the coexistence of D2D underlay communications and cellular communication causes serious co-channel interference, and an effective interference management is crucial. What makes the problem more complex, more and more articles confirm that the channel uncertainty cannot be ignored, especially in the high-speed mobile vehicular communication scenarios [17], [18]. In this paper, the Gauss-Markov process is proposed to statistically simulate the imperfect channel state information (CSI) [19], where the mobile characteristic of vehicles is highly considered. Furthermore, the chance constraint is used to describe the interference constraint in [20], which is in a probability form with uncertain parameters. To get the closed expression of the interference constraint, the authors proposed the Bernstein approximation method [21].

Motivated by combining the strengths of semantic communications and D2D technology, this paper proposes a long-term robust resource allocation scheme, in which joints access control of the application layer and robust power control of the physical layer. Through a series of optimization processes, this work is committed to realizing more efficient semantic information transmission.

A. Contributions

The main contributions of this work are shown as follows:

- The combination of semantic communication and a novel long-term resource allocation scheme is proposed in this paper to realize effective link transmission, which greatly improves transmission efficiency and resource utilization. The proposed scheme achieves the compromise of user satisfaction, queue stability, and communication delay on the premise of meeting user SINR requirements.
- Lyapunov optimization method is leveraged to transform dynamic cross-layer resource allocation problem into a semantic access control subproblem and a power control subproblem. Karush-Kuhn-Tucher (KKT) conditions and Lagrangian function method are used to deal with the subproblems, respectively. Particularly, the Bernstein approximation method is adopted to convert the non-convex power control subproblem into a solvable convex one.
- The mobility characteristics of vehicles are considered, and an accurate uncertain channel state description is constructed by introducing the first-order Markov process. A robust resource allocation algorithm is proposed to realize efficient and reliable semantic signal transmission.

The rest of this paper is organized as follows: In Section III, the system model and a robust resource allocation framework are established. In Section IV, the resource allocation problem transformation based on the Lyapunov optimization method is proposed. We propose a robust resource allocation algorithm in Section V. Numerical simulation results and conclusion are shown in Section VI and VII, respectively.

Notation: In this paper, vectors are typed by bold letters. Some notations shown in Table I.

TABLE I: Notations

$\mathcal{K}$	Index set of reused channels $\mathcal{K}=\{1,\ldots,K\}$
$\mathcal{M}$	User index set in a reused channel $\mathcal{M}$=$\{0,1,\ldots,M\}$
$\mathcal{T}$	Index set of time slots $\mathcal{T}=\{0,1,\ldots,T\}$
$\Pr\{\cdot\}$	Probability function
$E\{\cdot\}$	Exponential distribution
$\mathbb{E}\{\cdot\}$	Mathematical expectation
$\mathcal{R}^M$	Euclidean space
$\mathbb{R}^M$	Set of M-dimensional real vectors
$Q_m(t)$	Data queue of CUE and VUEs
$Z_m(t)$	Virtual queue of CUE and VUEs
$\mathbf{G}$	Channel gain vector
$\mathbf{D}$	Source data vector
$\mathbf{H}$	Encoded semantic data vector
$\mathbf{F}$	Received semantic data vector
$\hat{\mathbf{D}}$	Decoded semantic data vector

III. Problem Formulation

A. System Models

Based on Open System Interconnection (OSI) model, the cross-layer optimization diagram is shown in Fig. 1, where power control of the physical layer and access control of semantic data of the application layer are performed to realize queue stability. Taking the image information as an example, we apply the end-to-end scene graph generation model with

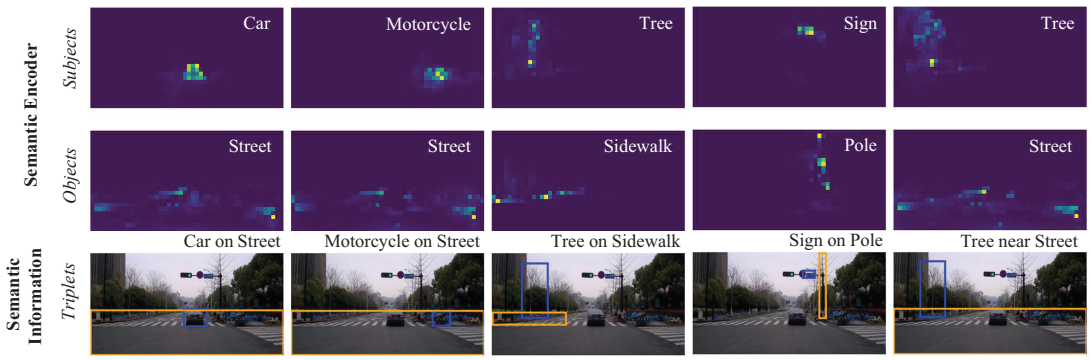

Fig. 2 Extracted image semantic data

encoder-decoder architecture to process the source image data [22]. In the semantic communication mechanism, the collected source data is represented as $\mathbf{D} = [d_1, d_2, \ldots]$, they are further encoded by semantic encoder and channel encoder. The extracted semantic data is represented by $\mathbf{H} = enc_c(enc_s(\mathbf{D}))$, where $enc_c(\cdot)$ and $enc_s(\cdot)$ are the channel encoder and the semantic encoder, respectively. At the end of receiving devices, the received signal is represented as $\mathbf{F} = \mathbf{GH} + \sigma^2$, where σ^2 is the background noise and $\mathbf{G}$ is the channel gain vector. Furthermore, The decoded semantic data is obtained by $\hat{\mathbf{D}} = dec_s(dec_c(\mathbf{F}))$, where $dec_c(\cdot)$ is the channel decoder, and $dec_s(\cdot)$ is the semantic decoder.

The extraction process of traffic semantic data is shown in Fig. 2. The D2D-V transmitter uses a semantic encoder to extract the semantic features from the real-time traffic images taken by vehicular cameras. These extracted key features are stored in a form of text and are cached in the buffer of the application layer. Based on these key features, vehicles can make rapid and efficient decisions. Meanwhile, the continuously generated semantic information can also be used to monitor real-time dynamic traffic environments and achieve safe and efficient autonomous driving. Furthermore, the semantic transmission process based on D2D communication technology is studied. The stability of the data queue is ensured by controlling the access rate and transmission rate. On the above basis, multiple indexes of system performance are optimized.

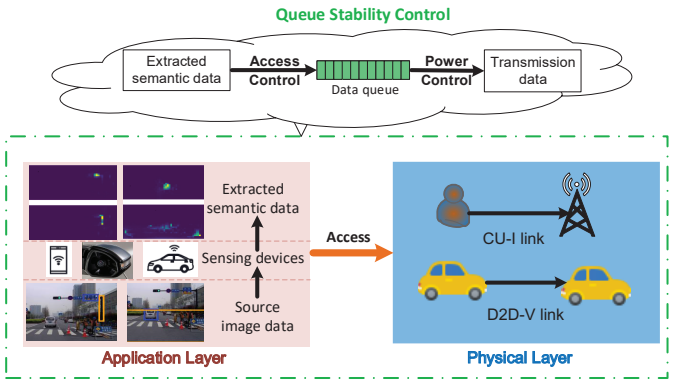

Fig. 1 Semantic communication-based cross-layer optimization model

Fig. 3 is the D2D-V communication model in the physical layer. Based on effective encoding and decoding technology, this paper focuses on the link transmission process. As shown in Fig. 3, a macrocell and numerous D2D-V pairs are included in the D2D-V networks, a vehicle transmitter (VT) and a vehicle receiver (VR) constitute a D2D-V pair. Each VT and VR is equipped with a semantic encoder and a semantic decoder, respectively. Multiple D2D-V pairs communicate directly by reusing the uplink allocated to CUE without passing through the base station (BS). When the distance between two neighbor vehicles exceeds the applicable distance of D2D communications, the D2D pairs are spontaneously formed. Cowan's M3 model can well describe the traffic pattern [23]. Cowan's M3 model stated that the distances between adjacent D2D pairs follow a truncated exponential distribution.

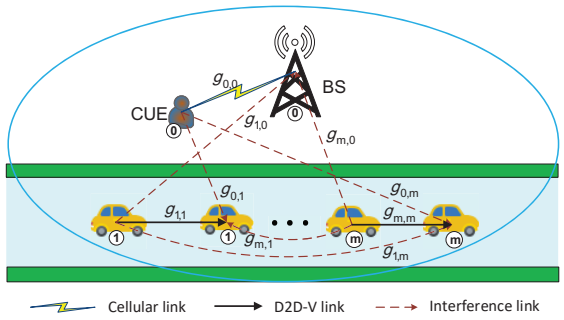

Fig. 3 Physical communication model

To improve the spectrum efficiency of semantic data transmission, a reusing mechanism is adopted where the CUE and VUEs are the spectrum owner and sharers, respectively. Fig. 3 also shows five kinds of links in a specific reused channel: CUE-I link between CUE and the BS, D2D-V link, V2I link between VT and the BS, CUE-V link between CUE and VR, V2V interference link between VT and VR. Particularly, the interference link and signal link are distinguished in this figure. The signal links include the CUE-I link and D2D-V link. V2I link, CUE-V link, and V2V interference link are the interference links. As a tricky feature of the vehicular communication scenarios, these links are regarded as Non-Line of Sight (NLoS) transmissions. It is believed that there are always many uncertain environmental factors in the semantic communication process, such as the obstruction of obstacles, the high-speed relative movement of communication terminals,

channel estimation error, and so on [24]. Therefore, channel uncertainty is inevitable for semantic data transmission, and the accurate description is shown in subsection B.

In the process of resource allocation, $M+1$ transmission users are considered (i.e., one CUE and M VTs), which are connected to the BS and M VRs, respectively. At the tth time slot, $A_m(t)$ bits of semantic data is required to transformed by user m, where $m \in \mathcal{M}$, $\mathcal{M} = \{0, 1, \ldots, M\}$. The arrival data is first stored in a buffer of the mth transmission user before being sent out. Let $R_m(p_m(t))$ denotes the transmission rate at time slot t of the physical layer, the queue m formed and is expressed as $Q_m(t)$. In other words, $A_m(t)$ and $R_m(p_m(t))$ specify the amount of semantic data that should arrive from the application layer and be sent out in the physical layer, respectively.

B. Channel Models

The channel gain between the mth transmitter and the nth receiver is formulated as

$$g_{m,n}^k = S_{m,n}^k (\eta_{m,n}^k)^2, \quad m \in \mathcal{M}, n \in \mathcal{M} \tag{1}$$

where $S_{m,n}^k$ and $(\eta_{m,n}^k)^2$ denote the large-scale fading and small-scal fading effects in the channel k, respectively [25], $k \in \mathcal{K}$, $\mathcal{K} = \{0, 1, \ldots, K\}$. The large-scale slow fading includes shadow fading and path loss,

$$S_{m,n}^k = L_{m,n}^k (d_{m,n}^k)^{-\alpha_m}, \quad m \in \mathcal{M}, n \in \mathcal{M} \tag{2}$$

where $L_{m,n}$ denotes the shadow fading and $d_{m,n}^{-\alpha_m}$ denotes the path loss, α_m is the path-loss exponent and $d_{m,n}$ is the communication distance.

The IoV is always in a dynamic communication scenario, vehicle movement will lead to the Doppler effect, thereby affecting the channel state. A well-function resource allocation strategy is inseparable from the capture of imperfect CSI. Therefore, the first-order Markov process is adopted to accurately simulate the small-scale fading component $\eta_{m,n}$, which is shown as follows [26]:

$$\eta = \vartheta \hat{\eta} + \epsilon, \tag{3}$$

where η and $\hat{\eta}$ are the channel responses of the current and previous time slots, respectively. The coefficient ϑ $(0 < \vartheta < 1)$ quantifies the channel correlation between the two consecutive time slots. In the probabilistic statistical model [26], ϵ is formulated by $\vartheta = J_0(2\pi f_d T_f)$, where $J_0(\cdot)$ is the zero-order Bessel function. T_f is the feedback time interval of the channel state information. $f_d = v f_c / c$ is the maximum Doppler frequency, where v is the relative speed between users, f_c denotes the carrier frequency, and $c = 3 \times 10^8$m/s. ϵ is the channel discrepancy term, which is independent to $\hat{\eta}$ and with the distribution of $\mathcal{CN}(0, 1-\vartheta^2)$.

The small-scale fading in the dynamic channel model is represented as follows,

$$(\eta_{m,n}^k)^2 = (\vartheta_{m,n}^k \hat{\eta}_{m,n}^k)^2 + (\epsilon_{m,n}^k)^2, \quad m \in \mathcal{M}, n \in \mathcal{M}. \tag{4}$$

The dynamic channel model is represented as follows,

$$g_{m,n}^k = S_{i,j}^k ((\vartheta_{m,n}^k \hat{\eta}_{m,n}^k)^2 + (\epsilon_{m,n}^k)^2), \quad m \in \mathcal{M}, n \in \mathcal{M}. \tag{5}$$

Given that $\hat{g}_{m,n}^k = S_{m,n}^k (\vartheta_{m,n}^k \hat{\eta}_{m,n}^k)^2$ and $\tilde{g}_{m,n}^k = S_{m,n}^k (\epsilon_{m,n}^k)^2$, (5) can be changed to

$$g_{m,n}^k = \hat{g}_{m,n}^k + \tilde{g}_{m,n}^k, \quad m \in \mathcal{M}, n \in \mathcal{M}, \tag{6}$$

where $\hat{g}_{m,n}^k$ denotes the sampling channel gain of the previous slots, $\tilde{g}_{m,n}^k$ denotes the error channel gain. $\hat{g}_{m,n}^k$ is a constant that can be measured, and $\tilde{g}_{m,n}^k \sim E(\frac{1}{S_{m,n}^k \left(1-(\vartheta_{m,n}^k)^2\right)})$ [25].

When $m = n = 0$, $g_{m,n}^k$ denotes the CUE-I link's channel gain in the kth channel; When $m = n \neq 0$, $g_{m,n}^k$ denotes the D2D-V links' channel gain; When $m \neq n$, $g_{m,n}^k$ denotes the interference links' channel gain, $g_{m,0}^k$, $g_{0,n}^k$, $g_{m,n}^k$ are the channel gains of V2I link, CUE-V link, V2V interference link in the kth uplink channel, respectively.

C. Cross-Layer Resource Allocation Problem Formulation

In this subsection, the cross-layer problem includes access control in the application layer and power allocation in the physical layer. Since the semantic data cannot be transmitted to VR instantaneously, the data has to be temporarily stored in the queue of the VT. In the time slot t, $t \in \mathcal{T}$, the queue backlogs of the mth VT are represented as

$$Q_m(t+1) = \max\{Q_m(t) - R_m(p_m(t)), 0\} + A_m(t), \tag{7}$$

It is shown that the dynamic data queue is composed of the transmission rate $R_m(p_m(t))$ and the access extracted semantic rate $A_m(t)$. The data transmission process is controlled by the power control strategy. The data access process is controlled by the semantic access rate control strategy. There is no data overflow if the transmission rate $R_m(p_m(t))$ is larger than or equal to the access rate $A_m(t)$ in the data queue $Q_m(t)$.

Definition 1: According to the definition of network stability, the data queue $Q_m(t)$ is mean rate stable [27] when

$$\lim_{T \to \infty} \frac{\mathbb{E}\{|Q_m(T)|\}}{T} = 0. \tag{8}$$

Channel reusing mechanism is assumed to improve spectrum efficiency in the physical layer. Under the limited spectrum resources, effective channel reusing is crucial for realizing wireless communication. However, the coexistence communications in the same frequency band will cause serious co-channel interference, the interference of the mth signal link is expressed as

$$I_m(t) = \sum_{n=0, n \neq m}^{M} p_n(t) g_{n,m}, \quad m \in \mathcal{M}, n \in \mathcal{M}, \tag{9}$$

where p_0 is the CUE's transmission power and I_0 is the interference of CUE-I link. When $n \geq 1$, p_n denotes the nth VT's power. Furthermore, the signal links' real-time SINR is formulated as

$$\gamma_m(p_m(t)) = \frac{p_m(t) g_{m,m}}{I_m(t) + \sigma^2}, \quad m \in \mathcal{M}, \tag{10}$$

The deterministic maximum equivalent transmission rate of VUEs calculated by Shannon's theorem is

$$R_m = \omega \log_2(1 + \bar{\gamma}_m(p_m(t))), \quad m \in \mathcal{M}. \tag{11}$$

where $\bar{\gamma}_m = \frac{\mathbb{E}\{p_m g_{m,m}\}}{\mathbb{E}\{\sum_{n=0,n\neq m}^{M} p_n g_{n,m}\}+\sigma^2} = \frac{p_m \bar{g}_{m,m}}{\sum_{n=0,n\neq m}^{M} p_n \bar{g}_{n,m}+\sigma^2}$.

The application-layer satisfaction U_m is positively related to the arrival rate. On the premise of maintaining the stability of the network, the throughput of the network can be expressed by the access rate of semantic date A_m. Therefore, user satisfaction U_m is positively related to the access rate in the application layer, and we defined U_m as a concave function, which is represented as

$$U_m[A_m(t)] = \Omega_m \log_2[A_m(t)], \tag{12}$$

where Ω_m is a predefined weight parameter of the mth user.

The transmission delay of the data queue at the mth user is defined as D_m. According to $Little'Law$, the average delay is represented as the quotient between the amount of access data and the transmission rate, which is shown as follows

$$\bar{D}_m(t) = \frac{\lim\limits_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{|Q_m(t)|\}}{\lim\limits_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{|R_m(p_m(t))|\}}. \tag{13}$$

The objective function is to optimize the long-term time-average satisfaction of CUE and VUEs. The cross-layer robust resource allocation problem is constructed as

$$\mathbf{P1}: \max_{A_m(t),p_m(t)} \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{\sum_{m=0}^{M} U_m[A_m(t)]\}$$
$$s.t. \begin{cases} C1: 0 \le p_m(t) \le p_{m,\max}, \quad \forall m,t \\ C2: 0 \le A_m(t) \le A_{m,\max}, \quad \forall m,t \\ C3: Q_m(t) \text{ is mean rate stable}, \quad \forall m,t \\ C4: \Pr\{\gamma_m(p_m(t)) \ge \gamma_{m,\min}\} \ge 1-\varepsilon, \ \forall m,t \\ C5: \bar{D}_m(t) \le D_{m,\max}, \ \forall m,t \end{cases} \tag{14}$$

where $C1$ is the power constraint, $p_{m,\max}$ is the maximal power. $C2$ is the constraint of access rate, and $A_{m,\max}$ is the maximum access rate of extracted semantic data. $C3$ represents the queue stability constraint defined in (8). $C4$ is the SINR constraint in the physical layer, $\gamma_{m,\min}$ is the SINR threshold, ε denotes the outage probability threshold of SINR constraint, where $\varepsilon \in (0,1)$. $C5$ represents the long-term constraint of the delay, and $D_{m,\max}$ is the delay threshold.

In the traditional static communication networks, we do not advocate expressing probability constraints of the user SINR in a short time slot by adopting a statistical model. However, in high-speed dynamic vehicular networks, the mobility characteristics may cause large distance changes, so it is necessary to constantly update the collected topology changes in the time slots. Since the adopted channel model includes path loss, while the communication distance is updated periodically, the statistical channel model parameters of the corresponding time slot should also be updated. Therefore, it is reasonable and necessary to use the short-term probability constraint $C4$ to describe the accurate service demands in mobile scenarios.

IV. Problem Transformation Based on Lyapunov Optimization

A. Problem Transformation

By exploiting the virtual queue concept [28], the long-term delay constraint in problem $\mathbf{P1}$ is converted to queue stability conditions. The virtual queue of the mth transmission user associated with delay constraint $C5$ is shown as follows

$$Z_m(t+1) = \max\{Z_m(t) - R_m(p_m(t))D_{m,\max}, 0\} + Q_m(t), \tag{15}$$

where the queue $Z_m(t)$ is not a real data queue, and (15) is just an equivalent queue that satisfies the constraint $C5$.

Theorem 1: If $Z_m(t)$ is mean rate stable, $C5$ holds automatically.

Proof: Since the space is limited, the process of the proof is omitted. The meticulous proof is shown in [29].

According to *Theorem 1*, problem $\mathbf{P1}$ is rewritten as

$$\mathbf{P2}: \max_{A_m(t),p_m(t)} \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{\sum_{m=0}^{M} U_m[A_m(t)]\}$$
$$s.t. \begin{cases} C1, C2, C4, \\ C6: Q_m(t), Z_m(t) \text{ are mean rate stable}, \ \forall m,t. \end{cases} \tag{16}$$

B. Lyapunov Optimization

Lyapunov optimization is a powerful method theory to deal with the long-term resource optimization schemes, which need less prior information and owns lower computational complexity [15]. Let $\mathbf{M}(t)=[\mathbf{Q}(t), \mathbf{Z}(t)]$ be the concatenated vector of the data queue and the virtual queue. Then, the Lyapunov function is defined as

$$L(\mathbf{M}(t)) = \frac{1}{2} \sum_{m=0}^{M} \{Q_m^2(t) + Z_m^2(t)\}. \tag{17}$$

The Lyapunov drift function is expressed as

$$\triangle(\mathbf{M}(t)) = \mathbb{E}\{L(\mathbf{M}(t+1)) - L(\mathbf{M}(t)) | \mathbf{M}(t)\}, \tag{18}$$

According to [30], a smaller drift value would have more conducive to queue stability. We can adjust the final queue length of the Lyapunov function to optimize the optimal value of Lyapunov drift, thereby realizing system stability. To minimize the network delay and maximize user satisfaction, the drift-minus-reward term is expressed as

$$\triangle(\mathbf{M}(t)) - V\mathbb{E}\{\sum_{m=0}^{M} U_m[A_m(t)]\}, \tag{19}$$

where V is a non-negative control parameter that affects the tradeoff between queue stability and user satisfaction.

Theorem 2: Define Δ_{max} as the upper bound of the drift-minus-reward, for all $\mathbf{M}(t)$ and $V \ge 0$, the maximum value of the drift-minus-reward term can be obtained by

$$\begin{aligned} \Delta_{max} = & \sum_{m=0}^{M} \mathbb{E}\{Q_m(t)A_m(t) - VU_m[A_m(t)] | \mathbf{M}(t)\} \\ & + \sum_{m=0}^{M} Z_m(t)\mathbb{E}\{Q_m(t) - R_m(p_m(t))D_{m,\max}^{Q} | \mathbf{M}(t)\} \\ & - \sum_{m=0}^{M} Q_m(t)\mathbb{E}\{R_m(p_m(t)) | \mathbf{M}(t)\} + \Theta, \end{aligned} \tag{20}$$

where Θ is a positive constant that satisfies the following constraint,

$$\begin{aligned} \Theta \ge & \frac{1}{2} \sum_{m=0}^{M} \{R_m^2(p_m(t)) + A_m^2(t) | \mathbf{M}(t)\} \\ & + \frac{1}{2} \sum_{m=0}^{M} \{(R_m(p_m(t))D_{m,\max}^{Q})^2 + Q_m^2(t) | \mathbf{M}(t)\}. \end{aligned} \tag{21}$$

Proof: The detailed proof is shown in Appendix A.

C. Joint Access Control and Power Control Optimization Scheme

Based on Lyapunov optimization theory, the rewritten objective function is regarded as the tradeoff between "network stability", "network delay", and "user satisfaction", and the optimization scheme should satisfy the constraints $C1$, $C2$, and $C4$. Therefore, the rewritten problem is

$$\begin{aligned} &\mathbf{P3}: \min \Delta_{max} \\ &s.t. \quad C1, C2, C4. \end{aligned} \tag{22}$$

Problem $\mathbf{P3}$ is divided into two independent subproblems, which are the access control subproblem and the power control subproblem.

1) Access Control Subproblem: The access rate control subproblem is expressed as

$$\begin{aligned} \mathbf{P4}: &\min_{A_m(t)} \sum_{m=0}^{M} Q_m(t)A_m(t) - VU_m[A_m(t)] \\ s.t. \quad &C2: 0 \le A_m(t) \le A_{m,\max}, \quad \forall m, t. \end{aligned} \tag{23}$$

2) Power Control Subproblem: The remaining items in the objective function are

$$\begin{aligned} &\sum_{m=0}^{M} Z_m(t)\Big(Q_m(t) - R_m(p_m(t))\Big)D^Q_{m,\max} \\ &- \sum_{m=0}^{M} Q_m(t)R_m(p_m(t)) + \Theta. \end{aligned} \tag{24}$$

Since the terms $Z_m(t)Q_m(t)$ and Θ involve no variables, the optimization objective function in the power control subproblem is formulated as

$$\sum_{m=0}^{M} \Big(Z_m(t)D^Q_{m,\max} + Q_m(t)\Big)R_m(p_m(t)). \tag{25}$$

Furthermore, the power control subproblem is given as

$$\begin{aligned} \mathbf{P5}: &\max_{p_m(t)} \sum_{m=0}^{M} \Big(Z_m(t)D^Q_{m,\max} + Q_m(t)\Big)R_m(p_m(t)) \\ s.t. &\begin{cases} C1: 0 \le p_m(t) \le p_{m,\max}, \quad \forall m, t \\ C4: \Pr\{\gamma_m(p_m(t)) \ge \gamma_{m,\min}\} \ge 1-\varepsilon, \ \forall m, t \end{cases} \end{aligned} \tag{26}$$

V. Solutions to Cross-Layer Resource Optimization Problem

A. Solution to Access Control Subproblem

KKT conditions is adopted to solve the optimization problem $\mathbf{P4}$. The Lagrangian function of problem $\mathbf{P4}$ is,

$$L_m(A_m(t)) = Q_m(t)A_m(t) - VU_m[A_m(t)] \tag{27}$$

where ν is the Lagrangian multiplier and $\nu \ge 0$.

The first-order derivative of (27) in terms of $A_m(t)$ is formulated as:

$$\frac{\partial L_m(A_m(t))}{\partial A_m(t)} = Q_m(t) - \frac{V\Omega_m}{A_m(t)\ln 2} = 0. \tag{28}$$

The optimal arrival rate of semantic data is obtained

$$A^*_m(t) = \min\{\tfrac{V\Omega_m}{Q_m(t)\ln 2}, A_{m,\max}\}. \tag{29}$$

B. Transformation of Power Control Subproblem

1) Successive Convex Approximation of the Objective Function: In the power control subproblem, the function is rewritten as

$$\sum_{m=0}^{M} (Z_m(t)D^Q_{m,\max} + Q_m(t))\omega \log_2(1+\bar{\gamma}_m(\mathbf{p}(t))). \tag{30}$$

The method of successive convex approximation is adopted to approximate the objective function by the theory of

$$\log_2(1+x) \ge \tfrac{1}{\ln 2}\left[X\ln(x) + Y\right], \tag{31}$$

where $x > 0$, X and Y are two coefficients that should be definitely settled.

Supposed that the equal form of lower bound approximation is acquired when $x = \bar{\gamma}_m(\mathbf{p}(t))$,

$$\log_2(1+\bar{\gamma}_m(\mathbf{p}(t))) = \tfrac{1}{\ln 2}\left[X_m \ln(\bar{\gamma}_m(\mathbf{p}(t))) + Y_m\right]. \tag{32}$$

According to the equality condition, (32) can be converted to,

$$(\tfrac{x}{\bar{\gamma}_m(\mathbf{p}(t))})^{X_m} \ge \tfrac{1+x}{1+\bar{\gamma}_m(\mathbf{p}(t))}. \tag{33}$$

For any X_m which meets (33), to obtain the lower-limit approximation, X_m is a valid coefficient for and less than 1. If $X_m \ge 1$, $(\frac{x}{\bar{\gamma}_m(\mathbf{p}(t))})^{X_m}$ is a concave function, there exists $x > 0$, which will cause that (33) is not valid. It can be learned that the function $y = \frac{1+x}{1+\bar{\gamma}_m(\mathbf{p}(t))}$ is a tangent line for $y = (\frac{x}{\bar{\gamma}_m(\mathbf{p}(t))})^{\frac{\bar{\gamma}_m(\mathbf{p}(t))}{1+\bar{\gamma}_m(\mathbf{p}(t))}}$ at $x = \bar{\gamma}_m(\mathbf{p}(t))$. Therefore, it is concluded that $X_m = \frac{\bar{\gamma}_m(\mathbf{p}(t))}{1+\bar{\gamma}_m(\mathbf{p}(t))}$ is the maximum value which satisfies (33), and $Y_m = \ln(1+\bar{\gamma}_m(\mathbf{p}(t))) - X_m\ln(\bar{\gamma}_m(\mathbf{p}(t)))$.

To acquire a standard convex objective function, the transformation $\tilde{p}_m(t) = \ln p_m(t)$ is introduced, and the standard convex optimization structure is obtained with respect to $\tilde{p}_m(t)$, and the prove process is shown in [31]. The lower bound of the objective function in problem $\mathbf{P5}$ is obtained

$$\begin{aligned} \max_{p_m(t)} \sum_{m=0}^{M} &\tfrac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max} + Q_m(t)) \\ &\cdot \left[X_m \ln(\bar{\gamma}_m(e^{\tilde{p}_m(t)})) + Y_m\right]. \end{aligned} \tag{34}$$

2) Approximation of Probability Constraint: To depress the uncertain probability constraint (10), the Bernstein approximation which is a convex approximation method is proposed [25]. The probability constraint is reformulated as

$$\Pr\left\{\phi_0(\mathbf{p}) + \sum_{m=0}^{M} \xi_m\phi_m(\mathbf{p}) \le 0\right\} \ge 1-\varepsilon, \tag{35}$$

where $\mathbf{p}$ is a deterministic variable vector, $\{\xi_m\}$ is a random variable with marginal distribution $\{\psi_m\}$. With the following conditions, inequality (35) is potentially satisfied for a given family of $\{\xi_m\}$ distributions,

1) $\{\phi_m(\mathbf{p})\}$ are affine in $\mathbf{p}$;
2) $\{\xi_m\}$ are independent of each other;
3) $\{\psi_m\}$ is with the bounded support of [-1,1], which is expressed as $-1 \le \psi_m \le 1, \forall m = 0, 1, \cdots, M$.

Theorem 3: By Bernstein approximation, the uncertain probability constraint $\Pr\{\gamma_m(p_m(t)) \geq \gamma_{m,\min}\} \geq 1-\varepsilon$ can be transformed into the l_∞-approximate constraint.

$$\sigma^2+\sum_{n=0}^{M}\chi_{n,m}e^{\tilde{p}_n(t)}+\sqrt{2\ln(\frac{1}{\varepsilon})}\sum_{n=0}^{M}\upsilon_{n,m}\leq 0. \tag{36}$$

$$\delta_{n,m}\beta_{n,m}e^{\tilde{p}_n(t)} \leq \sum_{n'=0}^{M}\upsilon_{n',m}. \tag{37}$$

Proof: The detailed proof is shown in Appendix B.

Based on *Theorem 3* and the approximated objective function (34), the power control subproblem $\mathbf{P5}$ can be transformed as,

$$\mathbf{P6}: \max_{\tilde{p}_m(t)} \sum_{m=0}^{M} \frac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t)) \cdot \left[X_m \ln(\bar{\gamma}_m(e^{\tilde{p}_m(t)}))+Y_m\right]$$

$$s.t. \begin{cases} \sigma^2+\sum\limits_{n=0}^{M}\chi_{n,m}e^{\tilde{p}_n(t)}+\sqrt{2M\ln(\frac{1}{\varepsilon})}\sum\limits_{n=0}^{M}\upsilon_{n,m}\leq 0 \\ \delta_{n,m}\beta_{n,m}e^{\tilde{p}_n(t)} \leq \sum\limits_{n'=0}^{M}\upsilon_{n',m} \\ -\infty \leq \tilde{p}_n(t) \leq \ln p_{m,\max}, \quad \forall m,t \end{cases} \tag{38}$$

where the problem $\mathbf{P6}$ is the l_∞-approximate power control subproblem.

C. Solution to the l_∞-approximate Power Control Subproblem

Since the problem $\mathbf{P6}$ is a standard convex problem, the Lagrangian function method is adopted to solve it.

Theorem 4: By solving the Lagrangian function of problem $\mathbf{P6}$, the iteration for the power control is formulated as

$$\begin{aligned} p_m^{l_\infty}(t+1) = \Big[&\ln\left(\frac{\omega X_m(t)}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\right) \\ &-\ln\Big(\frac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\sum_{n\neq m}^{M}X_n(t)\frac{\bar{\gamma}_n(e^{\tilde{\mathbf{P}}})\overline{g}_{m,n}}{e^{\tilde{p}_n}\overline{g}_{n,n}} \\ &+\sum_{n=0}^{M}(\zeta_m(t)\chi_{n,m}+\lambda_{n,m}(t)\sqrt{M}\delta_{n,m}\beta_{n,m})\Big)\Big]_{-\infty}^{\ln p_{m,max}}, \end{aligned} \tag{39}$$

where $[x]_i^j$=$\min\{\max\{x,i\},j\}$. ζ_m and $\lambda_{n,m}$ denote Lagrangian multipliers, $\mu_m \geq 0$ and $\lambda_{n,m} \geq 0$, which are shown

$$\begin{aligned} \lambda_{n,m}(t+1) = \Bigg[&\lambda_{n,m}(t)+ \\ &K_\lambda(t)(\sqrt{M}\delta_{n,m}\beta_{n,m}e^{\tilde{p}_n}+\frac{\sum_{m=0}^{M}\chi_{n,m}e^{\tilde{p}_n}-I_{th}}{\sqrt{-2\ln(\varepsilon)}})\Bigg]^+, \end{aligned} \tag{40}$$

$$\zeta_m(t+1) = \left(2\ln(\varepsilon)\right)^{-\frac{1}{2}}\sum_{n'=0}^{M}\mu_{n',m}(t+1), \tag{41}$$

where K_λ denotes the step-size.

Proof: See Appendix C

D. Robust resource allocation algorithm

We constructed a long-term cross-layer resource allocation problem (14) and proposed a robust resource allocation algorithm to solve it. Firstly, a series of values are set, including the maximum number of time slots T, the initial length of data queue $Q_m(0)$, the initial power $\tilde{p}_m(0)$, and the step size K. According to (29), we can obtain the optimal arrival rate of the application layer. Then, update the Lagrange multipliers $\lambda_{n,m}(t+1)$ of the l_∞-approximation. Furthermore, the power iteration expressions $p_m(t+1)$ are shown in (39). At last, the algorithm based on the Lyapunov optimization framework is shown as follows.

Algorithm 1 Robust resource allocation algorithm

1: Initialize
   - Set $T=100$, $t \leftarrow 1$.
   - Set $Q_m(0)=20$, $\tilde{p}_m(0)=-8$.
   - Set $K_\mu=0.1$, $K_\lambda=0.1$.
2: Initialize $\lambda_{n,m}>0$ for the l_∞-approximation.
3: **while** (A_m and p_m are not converged) and ($t<T$) **do**
4: **for** $\forall t \in \mathcal{T}$ **do**
5: Calculate the optimal arrival rate $A_m(t)$ by eq. (29).
6: Update $\lambda_{n,m}(t+1)$ and $\zeta_m(t+1)$ by eq. (40), (41).
7: Calculate the optimal power $p_m(t+1)$ by eq. (39).
8: **end for**
9: Set $t=t+1$.
10: **end while**

VI. Simulation And Performance Evaluation

Numerical simulations are shown here to verify the effectiveness of the robust resource allocation algorithm. First of all, the extraction process of traffic semantic data is finished and shown in Fig. 2 of Section III. In this process, the end-to-end scene graph generation model Relation Transformer for Scene Graph Generation (RelTR) is adopted, and we regard scene graph generation as a set prediction problem. The encoder reasons about the visual feature context and infers a set of fixed-size triples. Then, the extracted semantic data is transmitted by D2D communications. In the D2D-V networks, a simplified communication model involving one CUE and four D2D-V pairs is formulated, which is under the communication range of the BS. The corresponding parameters of the D2D-V system are shown in Table II.

Table II. System Parameters.

Parameters	Values
SINR threshold ($\gamma_{m,\min}$)	0.9
Delay threshold ($D_{m,\max}$)	0.1s
Bandwidth (ω)	10 MHz
Outage probability threshold (ε)	0.1
Background noise (δ^2)	-30 dBm
Control parameter (V)	75
Weight parameter related to the service (Ω_m)	1
Maximum power ($p_{m,\max}$)	0.02 W
Carrier frequency (f_c)	2 GHz
Feedback time interval (T_f)	2 ms
Speed of CUE	0 m/s
Speed of four D2D-V pairs	34, 30, 32, 30 m/s
Shadow fading $L_{m,n}$	0.5
Path-loss exponent α_m	2

Since the accessed semantic image data is still huge, access control and power control are combined to control the queue length and optimize the performance indicators in cross-layer optimization. Figures 4, 5, 6, 7 and 8 show the dynamic

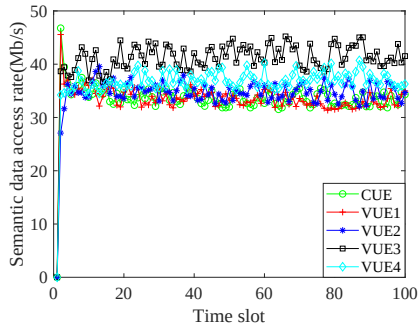

Fig. 4 Dynamic convergence of semantic access rates

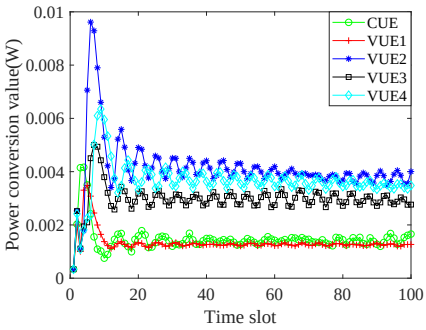

Fig. 5 Dynamic convergence of powers

convergence performance of the l_∞-approximation robust resource allocation algorithm in **Algorithm 1**. As depicted in Fig. 4 and 5, the access rates of semantic data and powers of CUE and VUEs all achieve dynamic convergence within several steps. Figures 6, 7, and 8 show the dynamic convergence performance of the data queue, the virtual queue, and the time delay, respectively. The backlogs of the data queue and the virtual queue also achieve the dynamic convergence within several steps, and the time delays of all users reach the ideal values which are less than the delay threshold $D_{m,\max}$. Therefore, the results in these figures demonstrate that the proposed robust resource allocation algorithm of the l_∞-approximation is effective and shows rapid convergence speed. It is noted that the long-term optimization scheme is different from traditional short-time ones, dynamic adjustment of power strategies is necessary to cope with uncertain channel changes, so the dynamic strategies contribute more robustness to the long-term semantic communication system.

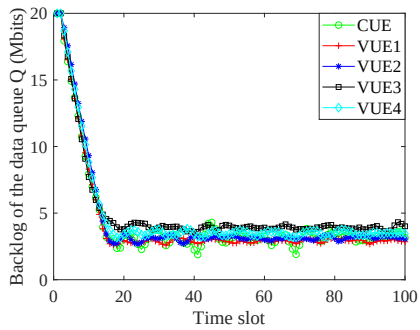

Fig. 6 Dynamic convergence of the data queue

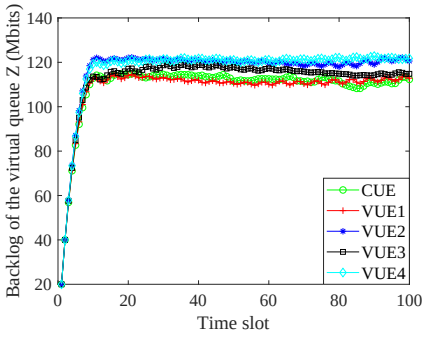

Fig. 7 Dynamic convergence of the virtual queue

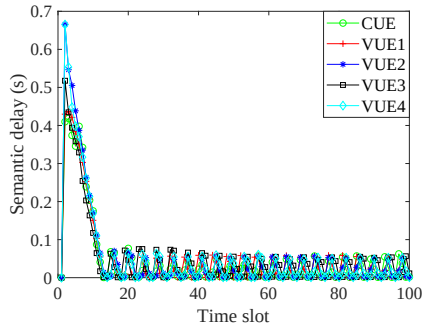

Fig. 8 Dynamic convergence of the time delay

Figures 9 and 10 illustrate the influence of the outage probability threshold ε on the long-term average sum transmission rate and time delay. According to Fig. 9, the sum transmission rate is higher when the outage probability threshold ε increases. A bigger value of ε means that the range of the variable p_m will be expanded, so the optimal power will be searched in a larger region and committed to improving the objective function. In this regard, the objective function is positively correlated with the access rate. When the system is stable, the input and output also reach a dynamic balance, so the increase of ε increases the access rate, and further increases the sum transmission rate. As shown in Fig. 10, the average time delay of CUE and four VUEs increases with the increase of ε. This is because the increase of ε brings a greater queue backlog, even if the transmission rate elevates, the average time delay will still increase. A comprehensive analysis of these two indicators in Fig. 9 and 10 can draw a conclusion that the transmission rate and delay restrict each other in the D2D-V communication system. The increase in transmission rate will lead to the loss of delay performance, which further reflects the importance of system performance compromising.

In this paper, the first-order Gauss-Markov process is used to describe the channel environment with imperfect CSI, and the l_∞-approximation robust resource allocation algorithm is proposed. To show the importance of considering channel uncertainty, we make a comparison with the benchmark [15], where perfect CSI is assumed. Besides, l_1-approximation, another method to transform the l_2-norm structure by $\|\mathbf{z}\|_2 \leq \|\mathbf{z}\|_1$ ($\mathbf{z} \in \mathbb{R}^M$) [25], is also simulated to compared with this

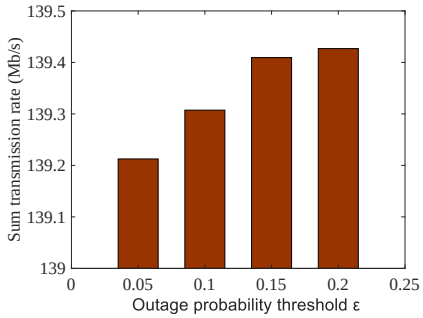

Fig. 9 Long-term sum transmission rate versus outage probability threshold ε

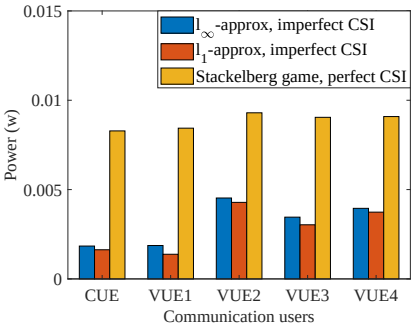

Fig. 12 Comparison of user powers in different cases

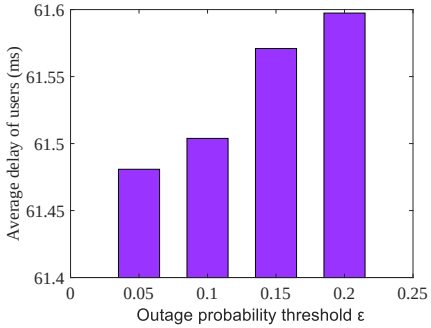

Fig. 10 Long-term average time delay versus outage probability threshold ε

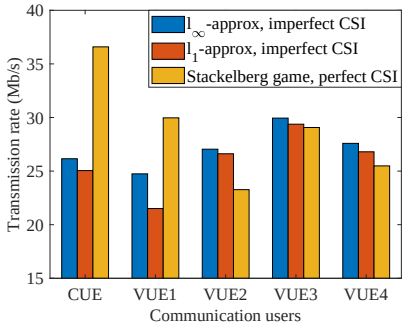

Fig. 13 Comparison of user transmission rates in different cases

work. Under the same target probability $\varepsilon = 0.1$, the comparison of the real outage probability is formulated. As shown in Fig. 11, the real outage probability of l_∞-approximation algorithm is lower than the l_1-approximation algorithm, and much lower than the benchmark [15]. According to $C4$ in the $\mathbf{P1}$, it is believed that the lower real outage probability means a better guarantee of the SINR constraint and the stronger system robustness. Therefore, this paper can achieve a more stable signal transmission than the benchmark [15] and the l_1-approximation method [25].

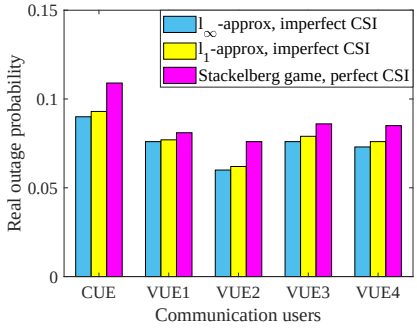

Fig. 11 Comparison of user real outage probability in different cases

The other performance indicators are also compared, and they are shown in Fig. 12, 13, and 14. It can be seen by comparing the l_∞-approximation and the l_1-approximation that the l_∞-approximation algorithm is outstanding in the performance of transmission rate and delay.

Moreover, by comparing with the l_∞-approximation algorithm, the three figures show that the benchmark [15] consumes much more transmission powers, but it does not get an obvious performance improvement. This is because the perfect CSI is impractical, and users are willing to increase the transmission rate by excessively increasing its powers. However, higher powers also bring more different-levels interference, and then the transmission rate and delay performances of different users will be affected to varying degrees. Therefore, the perfect CSI assumption always shows unstable and dissatisfied performances in the actual communication environment. To sum up, based on the comprehensive analysis as shown in Fig. 11, 12, 13, and 14, the l_∞-approximation algorithm is more well-function than the benchmark [15], especially for the system robustness and power consumption.

To verify the impact of vehicle mobility on sum transmission rate, different levels of vehicle speeds are simulated in D2D-V communication scenarios. In this numerical simulation, network topologies are assumed to be the same and the velocities of all VUEs are identical. Since the moving speed of CUE is slow, the CUE is assumed to be stationary. The unified

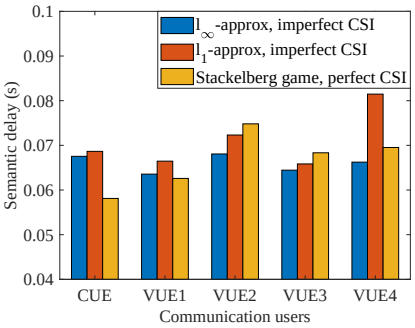

Fig. 14 Comparison of user delays in different cases

vehicle velocities are set to 0, 10, 20, and 30 m/s, respectively. As depicted in Fig. 15, when the vehicle speed increases, the sum transmission rate decreases in both the l_∞-approximation method of this paper and the l_1-approximation method of [25]. It is because that higher speed causes a more serious Doppler effect, which deteriorates the channel environment and makes the communication links suffer more co-channel interference. Therefore, taking the vehicle mobility characteristic into account is necessary for the accurate description of channel state and performance improvement.

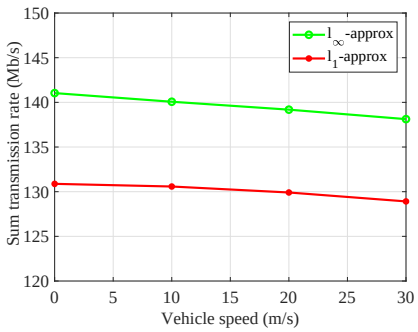

Fig. 15 Sum transmission rate versus vehicle speed in different cases

VII. Conclusion

In this paper, the extraction process of traffic semantic data is shown, and then the extracted semantic data is transmitted by D2D communications. Based on the Lyapunov optimization theory, this paper designs a long-term dynamic cross-layer resource allocation framework to realize effective semantic communications, which includes a semantic access control scheme and a power control scheme. A robust online resource allocation algorithm is proposed to achieve real-time optimization strategy. Simulation results demonstrated the converged performances of the proposed algorithms under the uncertain channel environment. The simulations also validate that the data queue is stabile, low delay and high-reliability semantic communications are finished, the user QoS requirements are guaranteed, especially the system robustness and power consumption outperform the benchmark. To sum up, the proposed long-term dynamic cross-layer resource allocation algorithm is well-function in the communication environment with multi-user interference and channel uncertainty. Moreover, the Lyapunov control parameter can be adjusted to realize the tradeoff between user satisfaction maximization, queue stability, and delay minimization.

Acknowledgement

This work is partly supported by National Natural Science Foundation of China under Grant 62273298, 61873223, the Graduate Innovation Foundation Projects of Hebei Province under grants CXZZBS2023055.

Appendix A

Proof: Assumed that A, B, and C are both non-negative real numbers, it is clearly established that

$$(\max\{A-B,0\}+C)^2 \le A^2+B^2+C^2+2A(C-B). \quad (42)$$

According to (42), we can obtain

$$\begin{aligned}&\triangle(\mathbf{M}(t)) = \mathbb{E}\{L(\mathbf{M}(t+1)) - L(\mathbf{M}(t))|\mathbf{M}(t)\}\\ &\le \sum_{m=0}^{M}\mathbb{E}\{Q_m(t)A_m(t)|\mathbf{M}(t)\}\\ &+ \sum_{m=0}^{M} Z_m(t)\mathbb{E}\{Q_m(t) - R_m(p_m(t))D^Q_{m,\max}|\mathbf{M}(t)\}\\ &- \sum_{m=0}^{M} Q_m(t)\mathbb{E}\{R_m(p_m(t))|\mathbf{M}(t)\} + \Theta\end{aligned} \quad (43)$$

where Θ is a positive constant that satisfies:

$$\begin{aligned}\Theta \ge \tfrac{1}{2}\sum_{m=0}^{M}\{R_m^2(p_m(t)) + A_m^2(t)|\mathbf{M}(t)\}\\ +\tfrac{1}{2}\sum_{m=0}^{M}\{(R_m(p_m(t))D^Q_{m,\max})^2 + Q_m^2(t)|\mathbf{M}(t)\}.\end{aligned} \quad (44)$$

Add $V\mathbb{E}\{\sum_{m=0}^{M} U_m[A_m(t)]\}$ to both sides of (43), we can obtain the formula (20). ■

Appendix B

Proof: The conservative approximation substitution for (35) is,

$$\inf_{\rho>0}\left[\phi_0(\mathbf{p}) + \rho\sum_{n=0}^{N}\iota_m(\rho^{-1}\phi_n(\mathbf{p})) + \rho\ln(\frac{1}{\varepsilon})\right] \le 0, \quad (45)$$

where $\iota_m(y) = \max_{\psi_m}\ln(\int\exp(xy)d\psi_m(x))$, ρ is the conservative approximate parameter. $\mathbf{p} = [p_1, p_2, \ldots, p_m]$, which is the vector of transmission powers.

The transformation process can further perform by using the upper bound of $\iota_m(y)$,

$$\iota_m(y) \le \max\{o_m^- y, o_m^+ y\} + \frac{\delta_m^2}{2}y^2, m = 0, 1, \cdots, M, \quad (46)$$

where o_m^-, o_m^+ and δ_m are both constants and determined by the given families $-1\le o_m^- \le o_m^+ \le 1$, $\delta_m \ge 0$.

When $\iota_m(\cdot)$ in (46) is substituted with the upper bound, the convex conservative surrogate of (45) is reformulated as,

$$\begin{aligned}\phi_0(\mathbf{p}(t)) + \sum_{m=0}^{M}\max\{o_m^-\phi_m(\mathbf{p}(t)), o_m^+\phi_m(\mathbf{p}(t))\}+\\ \sqrt{2\ln(\tfrac{1}{\varepsilon})}\left(\sum_{m=0}^{M}(\delta_m\phi_m(\mathbf{p}(t)))^2\right)^{\frac{1}{2}} \le 0.\end{aligned} \quad (47)$$

The outage probability constraint of the mth signal link in the optimization problem can be rewritten into a matrix expression,

$$\Pr\left\{(\mathbf{G}_m)^T\mathbf{p} + \sigma^2 \le 0\right\} \ge 1-\varepsilon, \quad (48)$$

where $\mathbf{G}_m = [g_{0,m}, g_{1,m}, \ldots, -\frac{g_{m,m}}{\gamma_{m,\min}}, \ldots, g_{M,m}]$. Here, $\tilde{g}_{n,m}$ is assumed to be bounded by $[a_{n,m}, b_{n,m}]$, $\beta_{m,n}=\frac{1}{2}(b_{m,n}-a_{m,n})\neq 0$ and $\varpi_{n,m}=\frac{1}{2}(b_{n,m}+a_{n,m})$ are constructed to normalize the support $\psi_{n,m}$, which is in the range of $[-1, 1]$.

$$\psi_{n,m} = \frac{\tilde{g}_{n,m} - \varpi_{n,m}}{\beta_{n,m}} \in [-1, 1]. \quad (49)$$

Let $\phi_0(\mathbf{p}(t))=\sigma^2+\sum_{i=0}^{N}(\hat{g}_{m,n}+\varpi_{m,n})p_m(t)$, $\phi_m(\mathbf{p}(t))=$

$\beta_{m,n}p_m(t)$, (47) is an equivalent constraint with $C4$ in $\mathbf{P5}$. Hence, substituting $f_0(\mathbf{p})$ and $f_m(\mathbf{p})$ into (47), then it is reformulated as,

$$\sigma^2+\sum_{n=0}^{M}\chi_{n,m}p_n(t)+\sqrt{2\ln(\tfrac{1}{\varepsilon})}\left(\sum_{n=0}^{M}(\delta_{n,m}\beta_{n,m}p_n(t))^2\right)^{\frac{1}{2}}\leq 0, \tag{50}$$

where $\chi_{n,m}=\hat{g}_{n,m}+\varpi_{n,m}+\upsilon_m^+\beta_{n,m}$. In (50), the coupling power variables bring a high computational difficulty to the problem solving process. To reduce the computational complexity, according to $\|\mathbf{z}\|_2 \leq \sqrt{N}\|\mathbf{z}\|_\infty$ ($\mathbf{z}\in\mathbb{R}^M$), the l_2-norm structure of (50) is further transformed into the l_∞-approximation problem,

$$\sigma^2+\sum_{n=0}^{M}\chi_{n,m}p_n(t)+\sqrt{2M\ln(\frac{1}{\varepsilon})}\max_{n\in\mathcal{M}}\delta_{n,m}\beta_{n,m}p_n(t)\leq 0 \tag{51}$$

By setting auxiliary variables $\upsilon=[\upsilon_{0,m},\upsilon_{1,m},\ldots,\upsilon_{M,m}]$, the l_∞-approximation constraint (51) can be further reformulated as the separable constraints (52) and (53).

$$\sigma^2+\sum_{n=0}^{M}\chi_{n,m}p_n(t)+\sqrt{2M\ln(\frac{1}{\varepsilon})}\sum_{n=0}^{M}\upsilon_{n,m}\leq 0 \tag{52}$$

$$\delta_{n,m}\beta_{n,m}p_n(t)\leq\sum_{n'=0}^{M}\upsilon_{n',m} \tag{53}$$

By the transformation $\tilde{p}_n(t)=\ln p_n(t)$, (36) and (37) can be obtained. ■

APPENDIX C

Proof: Since $\mathbf{P6}$ is a convex problem, the Lagrangian function of $\mathbf{P6}$ is shown as follows,

$$\begin{aligned}
&L(\tilde{p}_m(t):\lambda_{n,m},\zeta_m)=\\
&\sum_{m=0}^{M}\tfrac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\left[X_m\ln(\bar{\gamma}_m(e^{\tilde{p}_m(t)}))+Y_m\right]\\
&-\sum_{m=0}^{M}\zeta_m\left(\sigma^2+\sum_{n=0}^{M}\chi_{n,m}e^{\tilde{p}_m(t)}+\sqrt{2M\ln(\tfrac{1}{\varepsilon})}\sum_{n=0}^{M}\upsilon_{n,m}\right)\\
&-\sum_{m=0}^{M}\sum_{N=0}^{M}\lambda_{n,m}\left(\delta_{n,m}\beta_{n,m}e^{\tilde{p}_m(t)}-\sum_{n'=0}^{M}\upsilon_{n',m}\right),
\end{aligned} \tag{54}$$

where the lagrangian multipliers $\lambda_{n,m}\geq 0$ and $\zeta_m\geq 0$, respectively. Then, the corresponding dual function of the Lagrangian function is formulated as,

$$\begin{aligned}
&D(\lambda_{n,m},\zeta_m)=\max_{-\infty\leq\tilde{p}_i\leq\ln p_{i,max}}L(\tilde{p}_m(t):\lambda_{n,m},\zeta_m)\\
&=\max_{-\infty\leq\tilde{p}_i\leq\ln p_{i,max}}\sum_{m=0}^{M}\tfrac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\\
&\cdot\left[X_m\ln(\bar{\gamma}_m(e^{\tilde{p}_m(t)}))+Y_m\right]-\sum_{m=0}^{M}\zeta_m\sigma^2\\
&-\sum_{m=0}^{M}\sum_{n=0}^{M}\left(\zeta_m\chi_{n,m}+\lambda_{n,m}\sqrt{M}\delta_{n,m}\beta_{n,m}\right)e^{\tilde{p}_n}\\
&+\sum_{m=0}^{M}\sum_{n=0}^{M}\left(\sum_{n'=0}^{M}\lambda_{n',m}-\zeta_m\sqrt{2\ln(\tfrac{1}{\varepsilon})}\right)\upsilon_{n,m}.
\end{aligned} \tag{55}$$

Furthermore, the dual problem of (25) is as follows,

$$\min_{\lambda_{n,m}>0,\zeta_m>0}D(\lambda_{n,m},\zeta_m). \tag{56}$$

The power vector $\tilde{\mathbf{p}}$'s iteration function can be obtained by

$$\begin{aligned}
&\tfrac{\partial L(\tilde{p}_m(t):\lambda_{n,m},\zeta_m)}{\partial\tilde{p}_m}=\tfrac{\omega X_m}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\\
&-(\tfrac{\omega}{\ln 2}(Z_m(t)D^Q_{m,\max}+Q_m(t))\sum_{n\neq m}^{M}X_n\tfrac{\bar{\gamma}_n(e^{\tilde{\mathbf{p}}(t)})\overline{g}_{m,n}}{e^{\tilde{p}_n(t)}\overline{g}_{n,n}}\\
&+\sum_{n=0}^{M}\zeta_m^{(t)}\chi_{m,n}+\sum_{n=0}^{M}\mu_{m,n}^{(t)}(\sqrt{M}\sigma_{m,n}\alpha_{m,n})e^{\tilde{p}_m}=0,
\end{aligned} \tag{57}$$

where $\overline{g}_{m,n}$ and $\overline{g}_{n,n}$ are the expectation values of $g_{m,n}$ and $g_{n,n}$, respectively; $\overline{g}_{m,n}=\mathbb{E}\{g_{m,n}\}$ and $\overline{g}_{n,n}=\mathbb{E}\{g_{n,n}\}$, $m\in\mathcal{M}$ and $n\in\mathcal{M}$. Besides, $\bar{\gamma}_n(e^{\tilde{\mathbf{p}}(t)})$ is the average SINR of the nth signal link

$$\bar{\gamma}_n(e^{\tilde{\mathbf{p}}(t)})=\frac{e^{\tilde{p}_n(t)}\overline{g}_{n,n}}{\sum_{m\neq n}^{M}e^{\tilde{p}_m(t)}\overline{g}_{m,n}+\delta^2}. \tag{58}$$

Let $\frac{\partial L(\tilde{p}_m(t):\lambda_{n,m},\zeta_m)}{\partial\tilde{p}_m}=0$, the iteration for the power control is formulated as (39), and the iterations of the Lagrangian multipliers $\lambda_{n,m}$ and ζ_m are shown in (40) and (41), respectively. ■

REFERENCES

[1] F. Di, H. Christian, R. Lars, et al, "Deep multi-modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1341-1360, Mar. 2021.

[2] M. Xu, D. Hoang, J. Kang, et al, "Secure and reliable transfer learning framework for 6G-enabled Internet of Vehicles," *arXiv preprint*, arXiv:2111.05804, 2021.

[3] H. Zhang, N. Liu, X. Xiao, et al, "Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges," *IEEE Trans. Veh. Technol.*, vol. 55, no. 8, pp. 138-145, Aug. 2017.

[4] H. Xie and Z. Qin, "A lite distributed semantic communication system for internet of things," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 142-153, Jan. 2021.

[5] G. Shi, Y. Xiao, Y. Li, et al, "From semantic communication to semantic-aware networking: model, architecture, and open problems," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 44-50, Aug. 2021.

[6] Y. Shi, E. Alsusa, and M. Baidas, "Baidas joint DL/UL decoupled cell-association and resource allocation in D2D-underlay hetNets," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3640-3651, Apr. 2021.

[7] Z. Liu, J. Su, Y. Xie, et al, "Resource allocation in D2D enabled vehicular communications: A robust Stackelberg game approach based on price-penalty mechanism," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8186-8200, Jul. 2021.

[8] H. Xie, Z. Qin, G. Y. Li, and B. H. Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663-2675, Apr. 2021.

[9] Q. Gao, S. Lin, and G. Zhu, "Joint vehicular and static users multiplexing transmission with hierarchical modulation for throughput maximization in vehicular networks," *IEEE Intell. Transp. Syst. Mag.*, vol. 21, no. 9, pp. 3835-3847, Sept. 2020.

[10] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, no. 4, pp. 623-656, Oct. 1948.

[11] Y. Wang, M. Chen, W. Saad, et al, "Performance optimization for semantic communications: An attention-based learning approach," *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1-6.

[12] M. Yang and H. Kim, "Deep joint source-channel coding for wireless image transmission with adaptive rate control," *arXiv preprint*, arXiv.2110.04456, 2021.

[13] J. Mei, K. Zheng, L. Zhao, et al, "A latency and reliability guaranteed resource allocation scheme for LTE V2V communication systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3850-3860, Jun. 2018.

[14] S. Gong, P. Wang, and L. Duan, "Distributed power control with robust protection for pus in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3247-3258, Jun. 2015.

[15] Z. Zhou, Y. Guo, Y. He, et al, "Access control and resource allocation for M2M communications in industrial automation," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3093-3103, May. 2019.

[16] A. Asheralieva and D. Niyato, "Game theory and lyapunov optimization for cloud-based content delivery networks with device-to-device and UAV-enabled caching," *IEEE Trans. Veh. Technol.*, vol. 14, no. 3, pp. 1519-1532, Oct. 2015.

[17] Y. Xu, R. Hu, and G. Li, "Robust energy-efficient maximization for cognitive NOMA networks under channel uncertainties," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8318-8330, Sept. 2020.

[18] X. Li, L. Ma, Y. Xu, R. Shankaran, "Resource allocation for D2D-based V2X communication with imperfect CSI," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3545-3558, Apr. 2020.

[19] F. Fang, H. Zhang, J. Cheng, et al, "Joint User Scheduling and Power Allocation Optimization for Energy-Efficient NOMA Systems With Imperfect CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2874-2885, Dec. 2017.

[20] S. Wang, W. Shi, and C. Wang, "Energy-efficient resource management in OFDM-based cognitive radio networks under channel uncertainty," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3092-3102, Sept. 2015.

[21] A. Nemirovski and A. Shapiro, "Convex approximations of chance constrained programs," *SIAM J. Optim.*, vol. 17, no. 4, pp. 969-996, 2006.

[22] Y. Cong, M. Yang, and B. Rosenhahn, "RelTR: Relation Transformer for Scene Graph Generation," *arXiv preprint*, arXiv.2201.11460, 2022.

[23] L. Vasconcelos, A. Silva, . Seco, and J. Silva, "Estimating the parameters of Cowan's M3 headway distribution for roundabout capacity analyses," *Baltic J. Road Bridge Eng.*, vol. 7, no. 4, pp. 261-268, 2012.

[24] J. Pan, H. Shan, R. Li, et al, "Channel estimation based on deep learning in vehicle-to-everything environments," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 1654-1669, Jun. 2021.

[25] Z. Liu, Y. Xie, K. Y. Chan, et al. "Chance-constrained optimization in D2D-based vehicular communication network," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 5045-5058, May. 2019.

[26] T. Kim, D. Love, and B. Clerckx, "Does frequent low resolution feedback outperform infrequent high resolution feedback for multiple antenna beamforming systems?" *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1654-1669, Apr. 2011.

[27] S. Joshi, K. Manosha, M. Codreanu, and M. Latva, "Dynamic interoperator spectrum sharing via Lyapunov optimizations," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6365-6381, Oct. 2017.

[28] W. Bao, H. Chen, Y. Li, and B. Vucetic, "Joint rate control and power allocation for non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2798-2811, Dec. 2017.

[29] Y. Mao, J. Zhang, Y. Li, and K. Letaief, "A Lyapunov optimization approach for green cellular networks with hybrid energy supplies," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2463-2477, Dec. 2015.

[30] M. Neely, "Stochastic Network Optimization With Application to Communication and Queueing Systems," *San Rafael, CA, USA: Morgan and Claypool*, 2010.

[31] J. Papandriopoulos and J. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," *IEEE International Conf. on Commu.*, pp. 3270-3275, Dec. 2006.