Probabilistic Data Prefetching for Data Transportation in Smart Cities

Haichuan Ding, Ying Ma, *Student Member, IEEE*, Chi Zhang, *Member, IEEE*, Xuanheng Li, *Member, IEEE*, Bin Lin, *Senior Member, IEEE*, Yuguang Fang, *Fellow, IEEE*, and Shigang Chen, *Fellow, IEEE*

***Abstract*—To deal with the ever increasing wireless traffic, we have recently designed a vehicular cognitive capability harvesting network (V-CCHN) architecture to leverage vehicles as an alternative "transmission medium" (i.e., an opportunistic data carrier), besides the wireless spectrum, to effectively transport data from the location where it is collected to the place where it is consumed or utilized in a smart city environment. In the V-CCHN, cognitive radio technologies are utilized so that a large amount of data can be exchanged between vehicles and roadside infrastructure through short-range high-speed transmissions. Considering the limited contact duration and the uncertain activities of primary users, how to facilitate efficient data exchange between vehicles and roadside infrastructure is very challenging. This problem is further complicated by the fact that the mobility of vehicles might not be accurately predicted. In this paper, we propose a probabilistic data prefetching (PDP) scheme for the V-CCHN to address these challenges. By considering the conditional value at risk, we formulate the PDP schematic design as an optimization problem which allows us to obtain the corresponding PDP scheme. Finally, we have conducted extensive study to evaluate the performance of the obtained PDP scheme under various parameter settings.**

***Index Terms*—Conditional value at risk (CVaR), data offloading, data transportation, smart cities, vehicular networks.**

I. Introduction

THE RAPID development of information and communication technologies (ICTs) is revolutionizing the way we interact with our surrounding environment toward the era of smart cities [1]. As reported in [1]–[3], emerging mobile and smart-city applications will generate tremendous amount of wireless data which will further aggravate the congestion in radio access networks (RANs). Although network operators have been continuously investing on extra network resources, such as spectrum bands, they will soon be used up in handling the exploding wireless traffic. In consideration of recent initiatives on connected vehicles (CVs) and the change of traffic composition supported over telecommunications networks, we have recently proposed to leverage the mobility of vehicles in the cities as an alternative way, besides long-range direct wireless communications, to transport delay-tolerant data traffic from one place to another [4]–[7]. In our proposal, wireless communications is mainly used for short-range communications between corresponding transmitters/receivers and vehicles when they are getting close. Thanks to the short communication distance, the transmit power could be significantly reduced and the interference resulting from the communications between transmitters/receivers and vehicles will not significantly affect the utilization of spectrum resources. In so doing, we could potentially deliver large volume of delay-tolerant data traffic through the mobility of vehicles and save valuable wireless spectrum for delay-sensitive services. Along this line of thought, we design a novel emerging data transportation[1] network, the vehicular cognitive capability harvesting network (V-CCHN), where a secondary service provider (SSP) employs vehicles, equipped with cognitive radio routers (CR routers), as an alternative "transmission medium" (i.e., an opportunistic data carrier), besides wireless spectrum, to effectively transport data from the location where it is collected to the place where it is consumed or utilized in a smart city environment[2] [5]. For

[1]Data transportation refers to the process where data is carried on vehicles and delivered to intended places through the mobility of vehicles.

[2]Vehicles equipped with CR routers will be called CR router enabled vehicles (CRVs) in the subsequent development.

Manuscript received March 29, 2021; revised May 8, 2021; accepted May 26, 2021. Date of publication June 11, 2021; date of current version January 24, 2022. This work was supported in part by the U.S. National Science Foundation under Grant CNS-1717736, Grant CNS-1409797, and Grant CNS-1343356. The work of Xuanheng Li was supported in part by the National Natural Science Foundation of China under Grant 61801080; in part by the Project funded by China Postdoctoral Science Foundation under Grant 2020M680038; in part by the Fundamental Research Funds for the Central Universities under Grant DUT20RC(4)007; and in part by the Doctoral Research Initiation Fund of Liaoning Province under Grant 2019-BS-049. The work of Bin Lin was supported in part by the National Key Research and Development Program of China under Grant 2019YFE0111600, and in part by the LiaoNing Revitalization Talents under Program XLYC2002078. The work of Shigang Chen was supported in part by the U.S. National Science Foundation under Grant CNS-1909077. *(Corresponding author: Ying Ma.)*

Haichuan Ding is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: hcding@bit.edu.cn).

Ying Ma and Yuguang Fang are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: mayingbit2011@gmail.com; fang@ece.ufl.edu).

Chi Zhang is with the School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China (e-mail: chizhang@ustc.edu.cn).

Xuanheng Li is with the School of Information and Communication Engineering, Dalian University of Technology, Dalian 116023, China (e-mail: xhli@dlut.edu.cn).

Bin Lin is with the Information Science and Technology College, Dalian Maritime University, Dalian 116026, China, and also with the Network Communication Research Centre, Peng Cheng Laboratory, Shenzhen 518066, China (email: binlin@dlmu.edu.cn).

Shigang Chen is with the Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: sgchen@cise.ufl.edu).

Digital Object Identifier 10.1109/JIOT.2021.3088435

example, considering the nonuniform distribution of data traffic, the SSP can employ CRVs to pick up data from data networks in lightly loaded areas and carry data to the vicinity of mobile users. In this way, we can potentially relieve the congestion in heavily loaded areas [8]. To facilitate such services, the SSP should enable efficient data delivery from data networks to CRVs.

In this paper, we design a probabilistic data prefetching (PDP) scheme for the V-CCHN, where roadside infrastructure nodes exploit their predicted contact probabilities with CRVs to collaboratively prefetch the intended data for CRVs to ensure efficient data delivery from data networks to CRVs. Noticing that the amount of data to be prefetched is also related to the availability of spectrum resources, we cast the PDP schematic design under the V-CCHN architecture as a throughput maximization problem considering both uncertain activities of primary users (PUs) and the mobility of CRVs. In view of the inherent randomness in our design, to facilitate problem formulation, the conditional value at risk (CVaR) in finance and economics is introduced to quantify the amount of available spectrum resources and determine the importance of infrastructure nodes for each CRV when designing the data prefetching scheme. We demonstrate the effectiveness of the obtained PDP scheme by comparing it with the traditional approach where the same copy of data is prefetched to each of the infrastructure nodes that CRVs will meet in the future.

The remainder of this paper is organized as follows. The designed V-CCHN architecture and the basic idea of the PDP scheme are introduced in Section II. Based on the network model introduced in Section III, the PDP schematic design is cast as a throughput maximization problem in Section IV where PUs' uncertain activities and CRVs' uncertain mobility are handled with the CVaR. In Section V, we conduct performance evaluation to show the effectiveness of the proposed PDP scheme. Then, related work is reviewed in Section VI. Finally, conclusions are drawn in Section VII. To facilitate the easy reading, the important abbreviations and their definitions are summarized in Table I.

II. Network Architecture for the V-CCHN

As introduced in [5], the basic idea of the designed V-CCHN architecture is to employ partial fixed infrastructures and reliable bands to assist and supervise the utilization of dynamic networking resources, CRVs and harvested spectrum bands. Generally, the designed network architecture for the V-CCHN consists of an SSP, CR capable roadside units (CRSUs), and CRVs.

A. Major Network Entities

The SSP is a new emerging service entity who deploys partial physical infrastructure consisting of CRSUs and recruits CRVs, to jointly manage harvested spectrum resources to handle delay-tolerant data transportation and provide network services for various kinds of devices used in smart cities. An SSP can be an independent wireless service provider with certain amount of reliable spectrum bands called *basic bands*[3] or with leased bandwidth from existing wireless service providers (e.g., cellular service providers), which is willing to deploy partial infrastructure for new service provisioning. An SSP can be also an existing cellular service provider who wants to expand its advanced cellular services to users on the road or pick up newly added services for other CPSs or the Internet of Things (IoT). It is expected that the reliable spectrum bands (basic bands) are mainly utilized for common control signaling which provides fast information exchange between involved network entities in order to more effectively manage random variations in data traffic, harvested spectrum resources, and vehicular mobility.

TABLE I
LIST OF ABBREVIATIONS

Abbreviation	Definition
V-CCHN	Vehicular cognitive capability harvesting network
SSP	Secondary service provider
CR	Cognitive radio
CR router	Cognitive radio router
CRV	Vehicle equipped with a CR router
CRSU	CR capable roadside unit
SWB	Stationary wireless backhaul
PU	Primary user
PDP	Probabilistic data prefetching
LBR tuple	Link-band-radio tuple
CVaR	conditional value at risk

To facilitate data transportation, CRSUs with wired connections to data networks, called c-CRSUs, are placed at pivotal locations, such as major intersections, to aid network resource management, mobility prediction, and traffic management. These CRSUs allow the V-CCHN to gain access to backbone data networks. While other CRSUs that do not have wired connections to data networks, called r-CRSUs, can be deployed in the V-CCHN to form the stationary wireless backhaul (SWB) to help relay data traffic between CRVs and the backbone network. All CRSUs can exchange control messages with CRVs over reliable spectrum bands. The coverage of a c-CRSU's control signaling is called a *cell*. The logic architecture of a typical cell in the V-CCHN is illustrated in Fig. 1 where a c-CRSU acts as an agent for the SSP to oversee data transmissions and spectrum allocation within its corresponding cell. For ease of presentation, in the following, we say that the operations of CRVs and CRSUs under this cell are supervised by the SSP. Under the supervision of the SSP, CRVs and CRSUs under this cell work collectively and collaboratively to transport data to intended locations closer to end users or for data offloading.

To enable effective interactions with various systems/networks and efficient data transportation, both CRVs and CRSUs are equipped with CR routers as communication devices. CR routers are intelligent devices integrated with agile communication interfaces with CR and routing capabilities as well as abundant computing resources and

[3]Licensed spectrum bands, such as purchased stable TV bands or DSRC bands, with favorable propagation characteristics.

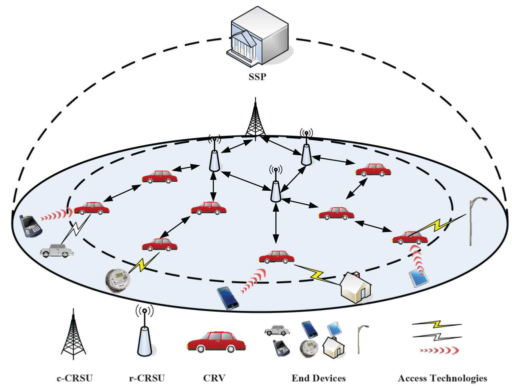

Fig. 1. Logic architecture of a typical cell in the V-CCHN.

storage spaces. CR routers can not only sense spectrum availability in a wide range, including TV white space and high frequency unlicensed bands, such as millimeter-wave (mmWave) bands, but also reconfigure their agile air interfaces into a wide spectrum of communications interfaces in order to interoperate with devices from various systems/networks. Moreover, they could exploit built-in computing resources and storage spaces to aggregate, store if needed, and carry data to the intended locations for delivery/offloading. Additionally, they are able to utilize computing resources to perform necessary edge computing functions to process data and user interactions, extract useful network intelligence, and help manage data transportation. Finally, the communication, computing, and storage capabilities of CR routers could be customized according to where they will be installed.

B. Data Transportation in the V-CCHN

Under the supervision of the SSP, CRVs carry data toward intended places and harvest spectrum resources for short-range high-speed vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) transmissions, for data transportation service provisioning. On the one hand, CRVs gather service requests/data from end devices and submit aggregated traffic requests to the SSP for coordination and resource allocation. On the other hand, CRVs serve as data mules for data delivery in the V-CCHN. Due to the constraints of road layouts, the mobility of CRVs is easier to predict. The SSP can take advantage of the predictability of the mobility of CRVs and employ them to forward data in the store-carry-forward manner. When necessary, the data-carrying CRVs exploit available harvested spectrum resources for short-range high-speed V2V or V2I transmissions to transfer data to other CRVs and CRSUs which can further carry/deliver data to the intended places. Once the SSP receives service requests from CRVs or other systems/networks, it makes network optimization to determine how and where the intended data should be delivered by jointly considering the types of the requested data as well as the availability of harvested spectrum bands and CRVs. After that, these decisions will be sent back to the corresponding c-CRSUs which manage spectrum resources and CRVs for data delivery within their corresponding cells. According to the management of c-CRSUs, CRVs transport data to designated locations either closer to end users or for offloading.

C. PDP Scheme

Depending on their locations, CRVs can obtain data from data networks either when they meet c-CRSUs or through the relaying of r-CRSUs. Notice that data is delivered from c-CRSUs to r-CRSUs through the relaying of the SWB by exploiting CR technologies and harvested spectrum resources. Thus, the achievable data rate between r-CRSUs and c-CRUSs is limited by the activities of PUs, which in turn limits the amount of data delivered to CRVs during their contacts with r-CRSUs.

In this paper, we consider addressing this challenge through data prefetching so that r-CRSUs can be effectively utilized to assist data exchange between CRVs and data networks [9]–[12]. Our basic idea is to let the SSP prefetch data to r-CRSUs which CRVs will meet in the future. In this way, a larger amount of data can be delivered to CRVs through r-CRSUs even the data rate between r-CRSUs and c-CRSUs is limited [9], [10].

In the V-CCHN, the SSP opportunistically exploits the mobility of CRVs for data delivery without controlling their trajectories. Although the mobility of CRVs is controlled by the physical constraints of roads, it is difficult for the SSP to exactly determine which r-CRSU a CRV will meet next [13]. In this case, an intuitive way to ensure data delivery is to prefetch the same copy of data to each of the r-CRSUs which the CRV will potentially meet in the future [9], [10]. This approach will not only incur more traffic in the SWB but also result in a waste of spectrum resources since the CRV will not meet all these r-CRSUs [11]. For illustrative purpose, let us consider the example shown in Fig. 2 where the SSP attempts to deliver data to CRV1. CRSU in Fig. 2 is a c-CRSU. CRSU1 and CRSU2 are r-CRSUs and one-hop neighbors. CRSU1/CRSU2 can receive data from CRSU through the relaying of the SWB. According to prediction, the SSP knows that CRV1 will turn right or go straight at the intersection with certain probabilities but cannot accurately predict the trajectory of CRV1. As a result, the SSP cannot exactly predict if CRV1 will meet CRSU1 or CRSU2. In this case, a simple approach for the SSP to ensure data delivery is to prefetch data to both CRSU1 and CRSU2 [12]. Since CRV1 will only meet either CRSU1 or CRSU2, the extra copy prefetched to the other r-CRSU will not be used and the corresponding spectrum resources used for prefetching this copy will be wasted, which leads to inefficient spectrum utilization in the SWB and calls for more efficient data prefetching schemes.

Clearly, the inefficiency of the aforementioned approach is resulted from the fact that the prefetched data is not fully utilized. If we could design a data prefetching scheme so that the prefetched data is utilized as much as possible, we could reduce the wasted spectrum resources and thus enable efficient spectrum utilization in the SWB for data prefetching. Intuitively, the traffic handled per r-CRSU at the edge of the SWB could be much less than that handled in the SWB. Thus, if the r-CRSUs that a CRV could meet in the future

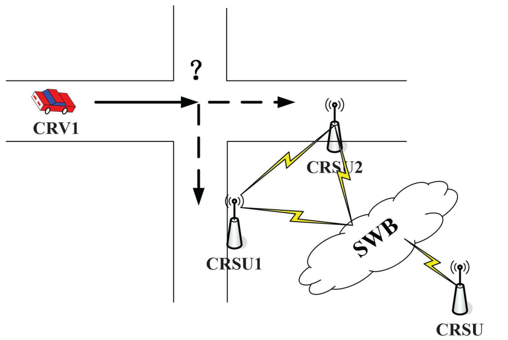

Fig. 2. Illustrative examples for data prefetching in the V-CCHN.

locate at the edge of the SWB, these r-CRSUs potentially have spare bandwidth to communicate with each other. In view of this, our basic idea to design such a data prefetching scheme is to exploit the available communication links between the involved r-CRSUs and let them collaboratively prefetch a copy of the intended data. Specifically, instead of prefetching a copy of the original data to each of the r-CRSUs, the SSP prefetches different parts of the data to each r-CRSUs. When the CRV meets one of these r-CRSUs, the SSP lets other r-CRSUs send their prefetched data to the corresponding r-CRSU. This can be illustrated through the example in Fig. 2. Rather than prefetching two copies of the original data to CRSU1 and CRSU2 via the SWB, the SSP can divide the data for CRV1 into two parts and push them to CRSU1 and CRSU2, respectively. Notice that CRSU1 and CRSU2 are one-hop neighbors. Once the SSP notices that CRV1 meets, for example, CRSU1, it will notify CRSU2 to send the obtained data content to CRSU1 via the communication link between them. In this way, both the data prefetched to CRSU1 and that prefetched to CRSU2 are used, which reduces the spectrum utilization in the SWB to prefetch the unused copies.

To fully exploit the benefit of the proposed scheme, the SSP should split data among the involved r-CRSUs so that the need for redirecting the prefetched data between r-CRSUs is kept to the minimum. In this paper, we attempt to design such a scheme, the PDP scheme, by exploiting the predicted contact probabilities between CRVs and r-CRSUs. In the PDP scheme, the SSP prefetches different parts of data, based on probabilistic contact information, to multiple r-CRSUs that a CRV will possibly meet next and employ the local communication links to adjust the prefetched data when the CRV meets one of these r-CRSUs.

Under the V-CCHN architecture, in the subsequent sections, we will study how to develop an effective PDP scheme to enable efficient interactions between CRVs and data networks by formulating an optimization problem and evaluate its performance accordingly.

III. Network Model

We consider a typical cell in the V-CCHN where N CRSUs are deployed by the SSP as partial infrastructure. One of these CRSUs is the c-CRSU, denoted as b, while other $N-1$ CRSUs are r-CRSUs. The r-CRSUs are indexed as $1, 2, \ldots, N-1$ and collectively form a SWB to assist data transmissions between CRVs and data networks. For simplicity, we collect all the CRSUs in a set $\Re = \{b, 1, \ldots, N-1\}$. The SSP tries to push data for K CRVs during a time period of T. To focus on the design of the PDP scheme, we assume that the SSP would like to prefetch data for these CRVs to the r-CRSUs which they will meet next. These CRVs are collected in a set $\mathcal{C}$. Due to uncontrollable mobility, the SSP only knows that the kth CRV will meet the nth r-CRSU with probability p_n^k. We collect the r-CRSUs with $p_n^k > 0$ in a set $\Re^k$. Namely, $\Re^k$ is the set of r-CRSUs which the kth CRV will potentially meet next. The set of available radios (i.e., transceivers for wireless data exchange) at i, $i \in \Re \cup \mathcal{C}$, is denoted as $\mathcal{H}_i$ and thus the number of radios available at i is the cardinality of $\mathcal{H}_i$, $|\mathcal{H}_i|$.

Considering the spatial variations in spectrum availability, the spectrum bands available to one CRSU might be different from that available to another CRSU. Mathematically, let $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$ be the set of spectrum bands, and $\mathcal{M}_i \subseteq \mathcal{M}$ represents the set of available bands at i, $i \in \Re \cup \mathcal{C}$. For $i \neq j$, $i, j \in \Re \cup \mathcal{C}$, $\mathcal{M}_i$ might be different from $\mathcal{M}_j$. Due to PUs' uncertain activities, the available time of these bands in $\mathcal{M}$ is subject to variations and thus the actual equivalent available bandwidth of these bands in the considered period of T might not be equal to the original bandwidth. In view of this, the equivalent available bandwidth of band m, $m \in \mathcal{M}_i \cap \mathcal{M}_j$, to link (i, j), $i, j \in \Re \cup \mathcal{C}$, is modeled as a random variable W_{ij}^m which is equal to the product of the original bandwidth of band m and the proportion of time where band m is available during the considered period.

The data transmission processes are characterized by the transmission ranges R_T and interference ranges R_I. The data transmission from the ith CRSU to the jth CRSU is assumed to be successful only if the jth CRSU is within the transmission range of the ith CRSU. We also assume that interference from the ith CRSU to the jth CRSU becomes nonnegligible if the jth CRSU is within the interference range of the ith CRSU. Noticing that the concepts of the transmission range and the interference range have been widely adopted in existing research works and their effectiveness has been confirmed in [14] and [15]. The transmission range and the interference range are closely related to the transmit power of CRSUs and the channel gain for wireless transmissions. Let us take the interference range as an example. The channel gain for wireless transmissions between CRSUs is modeled as

$$g_{ij} = \beta \cdot d_{ij}^{-\tilde{n}} \tag{1}$$

where β is an antenna related constant, $\tilde{n}$ is the path loss exponent, and d_{ij} is the distance between the ith CRSU and the jth CRSU. We assume that the transmission of CRSU i will cause nonnegligible interference to CRSU j if the power of the received signal exceeds a threshold of P_I. Then, the interference range of a CRSU can be derived from the aforementioned channel and interference models as

$$R_I = (\beta P / P_I)^{1/\tilde{n}} \tag{2}$$

where P is the transmit power of the CRSU.

In the following section, the PDP schematic design will be cast as a throughput maximization problem based on the models introduced in this section.

IV. Design of the PDP Scheme

In this section, we formulate the PDP schematic design as a throughput maximization problem. When making data prefetching decisions, the SSP needs to determine how much data could be pushed to individual r-CRSUs and reserve adequate resources so that CRVs can pick up the prefetched data from the corresponding r-CRSUs during contacts. When the SSP is not certain about the mobility of, for example, the kth CRV, it also needs to determine the amount of data to be prefetched to each r-CRSUs in $\Re^k$ and reserve enough resources for the potential data adjustment traffic generated among these r-CRSUs, based on the predicted contact probabilities. To efficiently schedule data transmissions for efficient data prefetching, it is necessary for the SSP to identify the conflicting relationships among different links. In the following section, we introduce a 3-D conflict graph to characterize the conflicting relationships in the V-CCHN. Then, we cast the PDP schematic design as a throughput maximization problem accordingly and employ CVaR to deal with the randomness in both W_{ij}^m and the mobility of CRVs.

A. 3-D Conflict Graph

The PDP scheme requires the SSP to not only allocate resources for prefetching data to r-CRSUs but also reserve enough resources for r-CRSUs to adjust prefetched data and for CRVs to obtain intended data from r-CRSUs. Clearly, such resource allocation problem will involve both CR links[4] in the SWB and CV links.[5] Thus, when making resource allocation decisions, the SSP needs to consider the potential conflicts between CR/CV links as well as those between CR links and CV links. In the considered V-CCHN, whether two communication links will conflict or not depends on not only the physical locations of their end devices but also the adopted bands and radios. Following [15], we view the considered V-CCHN as a 3-D resource space where dimensions are defined by links, available bands and radios. The basic resource unit in this resource space is a link-band-radio (LBR) tuple $((i, j), m, (h_i, h_j))$, where $i, j \in \Re \cup \mathcal{C}$, $m \in \mathcal{M}_i \cap \mathcal{M}_j$, $h_i \in \mathcal{H}_i$, and $h_j \in \mathcal{H}_j$. Noticing that the kth CRV will meet one of those r-CRSUs in $\Re^k$, we create a corresponding set of LBR tuples for the CV link from each r-CRSU in $\Re^k$ to the kth CRV. Then, the conflicts among CR/CV links in resource allocation can be characterized by the conflicts among these LBR tuples in the sense that conflicting LBR tuples cannot be scheduled at the same time.

The conflicting relationships among LBR tuples are captured via a 3-D conflict graph $G = (V, E)$ where each vertex $v \in V$ represents an LBR tuple and each edge $e \in E$ implies that the LBR tuples represented by its two end vertices cannot be scheduled at the same time. There exists an edge between

[4]The links between two CRSUs.

[5]The links from CRSUs to CRVs.

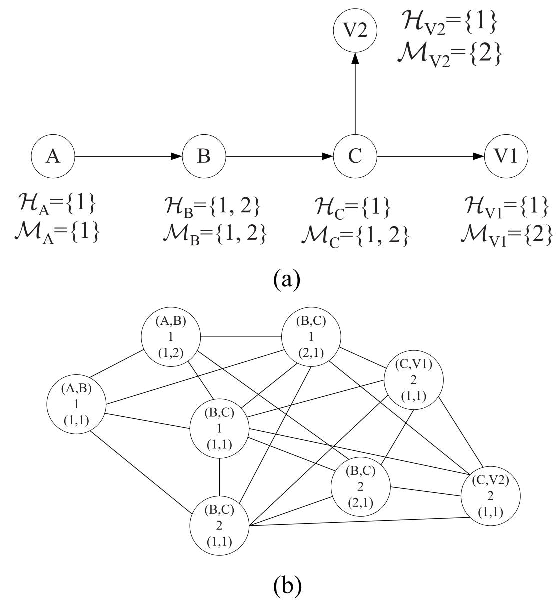

Fig. 3. 3-D conflict graph for a toy V-CCHN. (a) Toy topology of a V-CCHN. (b) 3-D conflict graph.

two vertices if and only if one of the following conditions is true for the corresponding two LBR tuples.

1) The links in these two LBR tuples are CR links. These two LBR tuples employ the same band and the receiving CRSUs in one tuple is within the interference range of the transmitting CRSUs in the other tuple.
2) The link in one LBR tuple is a CR link and the link in the other LBR tuple is a CV link. These two LBR tuples employ the same band and the receiving CRSUs in the tuple with CR link is within the interference range of the transmitting CRSUs in the other tuple with CV link.
3) If the links in these two LBR tuples are CV links, these two tuples will conflict when they share the same CRSU and band.
4) If the links in these two LBR tuples are CV links, these two tuples will conflict when they share the same CRSU and either have the same radios at the CRSU or share the same radio and the same CRV. In other cases, two LBR tuples conflict when they share the same radios at the same end devices.

The second condition follows from the fact that CV links are those links from CRSUs to CRVs. Considering short distances between CRSUs and CRVs, we assume that a CV link will only be interfered by another CV link which shares the same CRSU and spectrum resources, which leads to the second and the third conditions.[6] Since a CRV will meet only one CRSU in $\Re^k$, those LBR tuples with the same CRV will possibly conflict with each other only when they share the same CRSU. That is why we have the fourth condition.

For illustrative purpose, we consider the toy V-CCHN shown in Fig. 3(a) where A, B, C are CRSUs and $V1$

[6]If the SSP has more accurate characterization of the conflicting relationship between CV links, it can employ this information to refine the 3-D conflict graph and all the following formulations are still applicable.

and $V2$ are CRVs. A, B, C are assumed to have the same R_T and R_I with $R_T = 0.5R_I$. The distance between A and B as well as that between B and C equals R_T. The bands and radios available to A, B, C, $V1$, and $V2$ are shown in Fig. 3(a). Following the previous discussions, there exists an edge between vertex $((A, B), 1, (1, 1))$ and vertex $((B, C), 1, (2, 1))$ because of the first condition. There is an edge between vertex $((B, C), 2, (1, 1))$ and vertex $((C, V1), 2, (1, 1))$ because of both the second and the fourth conditions. Vertex $((C, V1), 2, (1, 1))$ conflicts with vertex $((C, V2), 2, (1, 1))$ due to the third and the fourth conditions. The conflicting relationships between other vertices can be similarly determined. The constructed conflict graph is shown in Fig. 3(b) where only the links indicated by arrows in Fig. 3(a) are considered for simplicity.

To maximize the spectrum utilization, the SSP should simultaneously schedule as many LBR tuples as possible. A set of LBR tuples can be scheduled at the same time when there exists no conflict between any two of these tuples, which implies that the set of vertices in G corresponding to these LBR tuples form an independent set of G. For an arbitrary independent set of G, if it becomes nonindependent when adding any one or more vertices, this independent set is called a maximal independent set (MIS) of G. To facilitate following problem formulation, we collect all the MISs of G in a set $\mathcal{I} = \{I_1, I_2, \dots, I_Q\}$, where I_q is the qth MIS of G and Q is the total number of MISs in $\mathcal{I}$. Next, the PDP schematic design will be formulated as a throughput maximization problem based on I_q's.

B. Flow Routing Constraints

In the considered V-CCHN, the SSP attempts to prefetch data for K CRVs. According to the PDP scheme, for the kth CRV, the SSP prefetches different parts of the intended data to r-CRSUs in $\Re^k$ and reserves resources in the local wireless links to adjust the prefetched data among different r-CRSUs when the kth CRV meets one of these r-CRSUs. As a result, the SSP should consider two kinds of flows when making resource allocation decisions. The flow generated to prefetch data to the lth r-CRSU in $\Re^k$ for the kth CRV is called the (k, l)th downloading flow and is denoted as $f^k(b, \xi_l)$, where b represents the c-CRSU with wired connection to data networks, k in the superscript means that the considered downloading flow is for the kth CRV, and ξ_l signifies that the destination of the considered flow (i.e., the lth r-CRSU in $\Re^k$). The flow generated to adjust prefetched data for the kth CRV is called the kth adjustment flow, denoted as $\widetilde{f}^k$. Different from downloading flows, $\widetilde{f}^k$ depends on which r-CRSU the kth CRV meets. As a result, the flow routing constraints for downloading flows and adjustment flows are presented separately.

Denote the rate of $f^k(b, \xi_l)$ as $\Upsilon_{k,l}$. The flow routing constraints for the (k, l)th downloading flow at the source b is

$$\sum_{j\in\mathcal{T}_b} f^k_{bj}(b, \xi_l) = \Upsilon_{k,l}$$
$$\sum_{\{i\in\Re-\{b\}|b\in\mathcal{T}_i\}} f^k_{ib}(b, \xi_l) = 0 \tag{3}$$

where $f^k_{bj}(b, \xi_l)$ is the rate of the (k, l)th downloading flow over link (b, j), $\mathcal{T}_b$ is the set of CRSUs within b's transmission range, and $\Re - \{b\}$ is the set of CRSUs without wired connections to data networks.

If the ith CRSU is an intermediate CRSU for $f^k(b, \xi_l)$, we have

$$\sum_{j\in\mathcal{T}_i} f^k_{ij}(b, \xi_l) = \sum_{\{j\in\Re-\{i\}|i\in\mathcal{T}_j\}} f^k_{ji}(b, \xi_l). \tag{4}$$

For ξ_l, the destination of $f^k(b, \xi_l)$, we have the following flow routing constraint:

$$\sum_{\{j\in\Re-\{\xi_l\}|\xi_l\in\mathcal{T}_j\}} f^k_{j\xi_l}(b, \xi_l) = \Upsilon_{k,l}. \tag{5}$$

Unlike $f^k(b, \xi_l)$, the sources and destinations of $\widetilde{f}^k$ are determined by which r-CRSU the kth CRV meets. If the kth CRV meets the nth r-CRSU in $\Re^k$, the SSP should direct those r-CRSUs in $\Re^k - \{\xi_n\}$ to deliver their received data for the kth CRV to the nth r-CRSU in $\Re^k$. For ease of presentation, we call the adjustment flow for the kth CRV between lth CRSU and the nth CRSU as the (k, l, n)th adjustment flow and denote it as $\widetilde{f}^k(\xi_l, \xi_n)$. Similar to downloading flows, we have the following set of constraints for $\widetilde{f}^k(\xi_l, \xi_n)$:

$$\sum_{j\in\mathcal{T}_{\xi_l}} \tilde{f}^k_{\xi_l j}(\xi_l, \xi_n) = \Upsilon_{k,l} \tag{6}$$

$$\sum_{\{j\in\Re-\{\xi_l\}|\xi_l\in\mathcal{T}_j\}} \tilde{f}^k_{j\xi_l}(\xi_l, \xi_n) = 0 \tag{7}$$

$$\sum_{j\in\mathcal{T}_i} \tilde{f}^k_{ij}(\xi_l, \xi_n) = \sum_{\{j\in\Re-\{i\}|i\in\mathcal{T}_j\}} \tilde{f}^k_{ji}(\xi_l, \xi_n) \tag{8}$$

$$\sum_{\{j\in\Re-\{\xi_n\}|\xi_n\in\mathcal{T}_j\}} \tilde{f}^k_{j\xi_n}(\xi_l, \xi_n) = \Upsilon_{k,l}. \tag{9}$$

By definition, $\widetilde{f}^k$ can be related to $\widetilde{f}^k(\xi_l, \xi_n)$ as

$$\tilde{f}^k_{ij} = \sum_{\xi_n\in\Re^k} \theta_{\xi_n} \sum_{\xi_l\in\Re^k-\{\xi_n\}} \tilde{f}^k_{ij}(\xi_l, \xi_n) \tag{10}$$

where $\theta_{\xi_n} \in \{0, 1\}$ indicates whether the kth CRV meets the nth CRSU in $\Re^k$. Considering the planned placement and limited coverage of CRSUs, we assume that each CRV connects to only one r-CRSU at each time point and thus have $\sum_{n=1}^{|\Re^k|} \theta_{\xi_n} = 1$, where $|\Re^k|$ is the cardinality of $\Re^k$.

C. Link Scheduling Constraints

The feasibility of aforementioned downloading flows and adjustment flows relies on the resource allocation in the data link layer. In this paper, we focus on time and MISs-based link scheduling. Namely, different MISs are scheduled to be active at different times. Denote the share of the considered time period T allocated to the qth MIS as λ_q ($\lambda_q \geq 0$) [16]. Then, we have

$$\sum_{q=1}^{Q} \lambda_q \leq 1, \quad \lambda_q \geq 0. \tag{11}$$

Since the data prefetched to r-CRSUs is intended for CRVs, the SSP should reserve enough resources at r-CRSUs so that

CRVs can pick up the prefetched data during contacts, which leads to the following constraint:

$$\theta_{\xi_n}\sum_{l=1}^{|\Re^k|}\Upsilon_{k,l} \le \sum_{q=1}^{Q}\lambda_q \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}} \frac{T_k}{T} W^m_{\xi_n\upsilon_k} r^m_{\xi_n\upsilon_k}\left(I_q\right) \quad (12)$$

where the left-hand side of (12) is the amount of data intended for the kth CRV, and the right-hand side represents the amount of resources should be reserved for the kth CRV at the nth r-CRSU in $\Re^k$. $W^m_{\xi_n\upsilon_k}$ is the equivalent available bandwidth between the nth r-CRSU in $\Re^k$ and the kth CRV over band m, T_k is the estimated length of the contact duration between the kth CRV and the nth r-CRSU in $\Re^k$ during the considered period T, and $r^m_{\xi_n\upsilon_k}(I_q)$ is defined as

$$r^m_{\xi_n\upsilon_k}\left(I_q\right) = \sum_{\substack{h_{\xi_n}\in\mathcal{H}_{\xi_n}\\ h_{\upsilon_k}\in\mathcal{H}_{\upsilon_k}}} r(\xi_n,\upsilon_k,m) \times 1\left(\left((\xi_n,\upsilon_k),m,\left(h_{\xi_n},h_{\upsilon_k}\right)\right)\in I_q\right) \quad (13)$$

where $r(\xi_n,\upsilon_k,m)$ is the maximum achievable data rate over 1 Hz bandwidth at link (ξ_n,υ_k) over band m. $1(\mathcal{A})$ is the indicator function that equals 1 when $\mathcal{A}$ is true and 0 otherwise.

Similarly, for links between CRSUs, the SSP should not only allocate resources for downloading flow but also reserve resources for adjustment flows. As a result, the corresponding link scheduling constraints are more complicated than existing works [17] and can be formulated as

$$\sum_{k=1}^{K}\left\{\sum_{\xi_l\in\Re^k} f^k_{ij}(b,\xi_l)+\tilde{f}^k_{ij}\right\} \le \sum_{q=1}^{Q}\lambda_q\sum_{m\in\mathcal{M}_i\cap\mathcal{M}_j} W^m_{ij} r^m_{ij}\left(I_q\right),\quad i,j\in\Re. \quad (14)$$

With (10), (14) can be reformulated as

$$\sum_{k=1}^{K}\sum_{\xi_l\in\Re^k}\left\{f^k_{ij}(b,\xi_l)+\theta_{\xi_l}\sum_{\xi_n\in\Re^k-\{\xi_l\}}\tilde{f}^k_{ij}(\xi_n,\xi_l)\right\} \le \sum_{q=1}^{Q}\lambda_q\sum_{m\in\mathcal{M}_i\cap\mathcal{M}_j} W^m_{ij} r^m_{ij}\left(I_q\right),\quad i,j\in\Re. \quad (15)$$

Due to the uncertainty in PUs' activities and CRVs' mobility, W^m_{ij}'s and θ_{ξ_l}'s in (12) and (15) are random variables, which makes (12) and (15) stochastic constraints. To proceed, we should find reasonable ways to quantify the values of W^m_{ij}'s and θ_{ξ_l}'s to reformulate (12) and (15) in order to deal with uncertainties. An intuitive approach is to replace W^m_{ij}'s and θ_{ξ_l}'s in (15) with their expected values. Unfortunately, this approach will lose necessary distribution information as shown in Fig. 4 where two random variables have the same expected value but with very different distributions. An alternative approach is to utilize the distribution information and quantify W^m_{ij}'s and θ_{ξ_l}'s with the values which can satisfy our purpose with high probability. From (15), θ_{ξ_l}'s are related to the amount of flow requested by the network layer and W^m_{ij}'s represent what each link can provide. In view of this, it is reasonable to quantify W^m_{ij}'s with values achievable with high probability and quantify θ_{ξ_l}'s with values not being exceeded with high probability. This observation motivates us to reformulate (12) and (15) based on the concept of the CVaR [18].

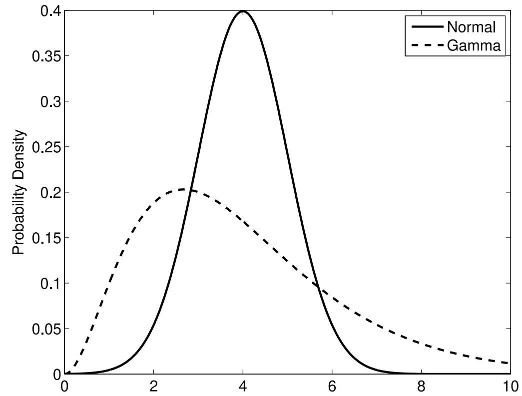

Fig. 4. Two random variables with the same expected value can have very different distributions. The normal random variable follows a distribution of $\mathcal{N}(4,1)$. The Gamma random variable follows a distribution of Gamma$(3, 0.75)$. Both of these two random variables have the expected value of 4.

1) Conditional Value at Risk: CVaR is a risk measure which provides us a way to quantify the values of random variables by considering their distributions [18], [19]. Due to its superior properties, CVaR has been widely used in finance and insurance industry for portfolio optimization to avoid large losses. For a random variable $Z=\underline{\omega}^T\underline{X}$, where $\underline{\omega}\in\mathbb{R}^{c\times 1}$ is a deterministic vector and $\underline{X}$ is a c dimensional random vector whose cumulative distribution function (CDF) is $\Phi_{\underline{X}}(\underline{x})$, its CVaR is defined as

$$\chi^\alpha_Z\left(\underline{\omega}\right) = \mathrm{E}_{\Phi^\alpha_Z(z)}[Z] \quad (16)$$

where the superscript α in $\chi^\alpha_Z(\underline{\omega})$ represents the confidence level that the value of Z is less than or equal to $z_\alpha(\underline{\omega})$, $z_\alpha(\underline{\omega}) = \min\{z|\Phi_Z(z)\ge\alpha.\}$, where $\underline{\omega}$ in $\chi^\alpha_Z(\underline{\omega})$ and $z_\alpha(\underline{\omega})$ is to signify their dependence on the vector $\underline{\omega}$. $\Phi^\alpha_Z(z)$ is the conditional α-tail distribution defined as

$$\Phi^\alpha_Z(z) = \begin{cases} \frac{\Phi_Z(z)-\alpha}{1-\alpha}, & z\ge z_\alpha\left(\underline{\omega}\right)\\ 0, & z< z_\alpha\left(\underline{\omega}\right)\end{cases} \quad (17)$$

where $\Phi_Z(z)$ is the distribution function of Z. From (16) and (17), CVaR is actually the mean of the α-tail distribution of Z. Thus, by setting the value of α and quantify Z with $\chi^\alpha_Z(\underline{\omega})$, we can guarantee that Z takes a value less than or equal to its CVaR with probability at least α.

$\chi^\alpha_Z(\underline{\omega})$ can be expressed through a minimization formula as [18]

$$\chi^\alpha_Z\left(\underline{\omega}\right) = \min_{\zeta\in\mathbb{R}}\zeta+\frac{1}{1-\alpha}\mathrm{E}\left\{[Z-\zeta]^+\right\} = \min_{\zeta\in\mathbb{R}}\zeta+\frac{1}{1-\alpha}\mathrm{E}\left\{\left[\underline{\omega}^T\underline{X}-\zeta\right]^+\right\} \quad (18)$$

where $[.]^+=\max\{0,.\}$. From [18], $\chi^\alpha_Z(\underline{\omega})$ has two very useful properties as shown in the following Lemmas, where we provide alternative proofs of these two properties based on (18) for completeness.

Lemma 1: $\chi_Z^\alpha(\underline{\omega})$ is positively homogeneous in $\underline{\omega}$, i.e.,

$$\chi_Z^\alpha(a\underline{\omega}) = a\chi_Z^\alpha(\underline{\omega}) \quad \forall a > 0. \tag{19}$$

Proof: From (18), we have

$$\begin{aligned}\chi_Z^\alpha(a\underline{\omega}) &= \min_{\zeta\in\mathbb{R}} \zeta + \frac{1}{1-\alpha}\mathrm{E}\left\{\left[a\underline{\omega}^T\underline{X} - \zeta\right]^+\right\}\\ &= a\min_{\zeta\in\mathbb{R}} \frac{\zeta}{a} + \frac{1}{1-\alpha}\mathrm{E}\left\{\left[\underline{\omega}^T\underline{X} - \frac{\zeta}{a}\right]^+\right\}\\ &= a\chi_Z^\alpha(\underline{\omega}).\end{aligned} \tag{20}$$

■

Lemma 2: $\chi_Z^\alpha(\underline{\omega})$ is a convex function of $\underline{\omega}$, i.e.,

$$\begin{aligned}&\chi_Z^\alpha(a_1\underline{\omega}_1 + a_2\underline{\omega}_2) \le a_1\chi_Z^\alpha(\underline{\omega}_1) + a_2\chi_Z^\alpha(\underline{\omega}_2),\\ &\quad \forall 0 \le a_1, a_2 \le 1, a_1 + a_2 = 1.\end{aligned} \tag{21}$$

Proof: According to (18), we have

$$\begin{aligned}&\chi_Z^\alpha(a_1\underline{\omega}_1 + a_2\underline{\omega}_2)\\ &= \min_{\zeta\in\mathbb{R}} \zeta + \frac{1}{1-\alpha}\mathrm{E}\left\{\left[a_1\underline{\omega}_1{}^T\underline{X} + a_2\underline{\omega}_2{}^T\underline{X} - \zeta\right]^+\right\}.\end{aligned} \tag{22}$$

Denote ζ_1^* and ζ_2^* as

$$\begin{aligned}\zeta_1^* &= \arg\min_{\zeta\in\mathbb{R}}\left\{\zeta + \frac{1}{1-\alpha}\mathrm{E}\left\{\left[\underline{\omega}_1{}^T\underline{X} - \zeta\right]^+\right\}\right\}\\ \zeta_2^* &= \arg\min_{\zeta\in\mathbb{R}}\left\{\zeta + \frac{1}{1-\alpha}\mathrm{E}\left\{\left[\underline{\omega}_2{}^T\underline{X} - \zeta\right]^+\right\}\right\}.\end{aligned} \tag{23}$$

Together with (22), we have

$$\begin{aligned}&\chi_Z^\alpha(a_1\underline{\omega}_1 + a_2\underline{\omega}_2)\\ &\le (a_1\zeta_1^* + a_2\zeta_2^*) + \frac{1}{1-\alpha}\\ &\quad\times \mathrm{E}\left\{\left[a_1\underline{\omega}_1{}^T\underline{X} + a_2\underline{\omega}_2{}^T\underline{X} - (a_1\zeta_1^* + a_2\zeta_2^*)\right]^+\right\}\\ &\le a_1\zeta_1^* + \frac{1}{1-\alpha}\mathrm{E}\left\{a_1\left[\underline{\omega}_1{}^T\underline{X} - \zeta_1^*\right]^+\right\}\\ &\quad + a_2\zeta_2^* + \frac{1}{1-\alpha}\mathrm{E}\left\{a_2\left[\underline{\omega}_2{}^T\underline{X} - \zeta_2^*\right]^+\right\}\\ &= a_1\chi_Z^\alpha(\underline{\omega}_1) + a_2\chi_Z^\alpha(\underline{\omega}_2).\end{aligned} \tag{24}$$

■

2) Constraint Reformulation: Clearly from above discussions, θ_{ξ_l}'s can be quantified with $\chi_{\theta_{\xi_l}}^\alpha(1)$'s. For W_{ij}^m's, we quantify it with $-\chi_{-W_{ij}^m}^\alpha(1)$ and the rationale for this choice is shown in the following theorem.

Theorem 1: $-\chi_{-W_{ij}^m}^\alpha(1)$ is achievable for W_{ij}^m with the probability of at least α.

Proof: By definition, we have

$$\begin{aligned}\mathrm{P}\left(W_{ij}^m \ge -\chi_{-W_{ij}^m}^\alpha(1)\right) &= \mathrm{P}\left(-W_{ij}^m \le \chi_{-W_{ij}^m}^\alpha(1)\right)\\ &= \Phi_{-W_{ij}^m}\left(\chi_{-W_{ij}^m}^\alpha(1)\right) \ge \alpha.\end{aligned} \tag{25}$$

■

Then, we could reformulate (12) and (15) with CVaR as

$$\begin{aligned}&-\sum_{q=1}^{Q}\lambda_q \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}} \frac{T_k}{T}\chi_{-W_{\xi_n\upsilon_k}^m}^\alpha(1) r_{\xi_n\upsilon_k}^m(I_q)\\ &\ge \chi_{\theta_{\xi_n}}^\alpha(1)\sum_{l=1}^{|\Re^k|}\Upsilon_{k,l}, n, l \in \Re^k.\end{aligned} \tag{26}$$

$$\begin{aligned}&\sum_{k=1}^{K}\sum_{l\in\Re^k}\left\{f_{ij}^k(b,l) + \chi_{\theta_l}^\alpha(1)\sum_{n\in\Re^k-\{l\}}\tilde{f}_{ij}^k(n,l)\right\}\\ &\le -\sum_{q=1}^{Q}\lambda_q \sum_{m\in\mathcal{M}_i\cap\mathcal{M}_j}\chi_{-W_{ij}^m}^\alpha(1) r_{ij}^m(I_q).\end{aligned} \tag{27}$$

With (26) and (27), we can provide an additional rationale for employing CVaR to quantify θ_{ξ_l}'s and W_{ij}^m's, which is shown in Theorem 2.

Theorem 2: With (26) and (27), it can be guaranteed that (12) and (15) are valid with the probability of at least α.

Proof: We will prove the results for (26), and the results for (27) can be proved in the same way. For simplicity, denote $\sum_{l=1}^{|\Re^k|}\Upsilon_{k,l}$ as ℓ_0. Notice that the lefthand side of (26) can be reformulated as a linear combination of $\chi_{-W_{\xi_n\upsilon_k}^m}^\alpha(1)$'s with coefficients $\ell_{\xi_n\upsilon_k}^m = \sum_{q=1}^{Q}\lambda_q(T_k/T)r_{\xi_n\upsilon_k}^m(I_q)$. Then, (26) can be reformulated as

$$\ell_0\chi_{\theta_{\xi_n}}^\alpha(1) + \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}}\ell_{\xi_n\upsilon_k}^m\chi_{-W_{\xi_n\upsilon_k}^m}^\alpha(1) \le 0. \tag{28}$$

By definition of CVaR, we have

$$\chi_{\theta_{\xi_n}}^\alpha(1) = \chi_Z^\alpha\left(\underline{e}_{\theta_{\xi_n}}\right) \tag{29}$$

$$\chi_{-W_{\xi_n\upsilon_k}^m}^\alpha(1) = \chi_Z^\alpha\left(\underline{e}_{\xi_n\upsilon_k}^m\right) \tag{30}$$

where $Z = \underline{\omega}^T\underline{X}$, $\underline{\omega}$ is the weight vector, $\underline{X}$ is a vector consisting of θ_{ξ_n} and $-W_{\xi_n\upsilon_k}^m$, $\underline{e}_{\theta_{\xi_n}}$ is a vector whose element corresponding to θ_{ξ_n} is 1 and other elements equal 0, $\underline{e}_{\xi_n\upsilon_k}^m$ is a vector whose element corresponding to $-W_{\xi_n\upsilon_k}^m$ is 1 and other elements equal 0. Based on Lemma 1 and Lemma 2, we have

$$\begin{aligned}&\ell_0\chi_{\theta_{\xi_n}}^\alpha(1) + \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}}\ell_{\xi_n\upsilon_k}^m\chi_{-W_{\xi_n\upsilon_k}^m}^\alpha(1)\\ &= \ell_0\chi_Z^\alpha\left(\underline{e}_{\theta_{\xi_n}}\right) + \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}}\ell_{\xi_n\upsilon_k}^m\chi_Z^\alpha\left(\underline{e}_{\xi_n\upsilon_k}^m\right)\\ &\ge \chi_Z^\alpha\left(\ell_0\underline{e}_{\theta_n} + \sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}}\ell_{\xi_n\upsilon_k}^m\underline{e}_{\xi_n\upsilon_k}^m\right) = \chi_Z^\alpha(\underline{\ell})\end{aligned} \tag{31}$$

where $\underline{\ell}$ is a vector formed by ℓ_0 and $\ell_{\xi_n\upsilon_k}^m$'s. Combining (28) and (31), we have $\chi_Z^\alpha(\underline{\ell}) \le 0$. Following the definition of CVaR, this implies that:

$$\mathrm{P}\left(\underline{\ell}^T\underline{X} \le 0\right) \ge \alpha. \tag{32}$$

Reformulating (32) based on the definitions of $\underline{\ell}^T$ and $\underline{X}$, it follows:

$$\mathrm{P}\left(\theta_{\xi_n}\sum_{l=1}^{|\Re^k|}\Upsilon_{k,l} \le \sum_{q=1}^{Q}\lambda_q\sum_{m\in\mathcal{M}_{\xi_n}\cap\mathcal{M}_{\upsilon_k}}\frac{T_k}{T}W_{\xi_n\upsilon_k}^m r_{\xi_n\upsilon_k}^m(I_q)\right) \ge \alpha. \tag{33}$$

Clearly from (33), the constraint in (12) is satisfied with the probability of at least α when (26) is satisfied. Following the same procedure, we could prove the same result for (15) and (27). Theorem 2 is proved. ■

Thus, we will use (26) and (27) instead of (12) and (15) as our link scheduling constraints.

D. Efficient Data Prefetching Under Multiple Constraints

To efficiently utilize CRVs for service provisioning, the SSP attempts to deliver as much data to CRVs as possible. Namely, the SSP prefetches as much data to those r-CRSUs in $\Re^k$, $k = 1, \ldots, K$ as possible. As aforementioned, in the PDP scheme, the SSP should not only allocate resources for the downloading flows but also reserve resources for the adjustment flows and the CV links so that CRVs are able to obtain the intended data contents. In view of this, when studying the PDP scheme design, we should jointly consider the resources demanded by the downloading flows, the adjustment flows, and the CV links. With above flow routing and link scheduling constraints, we formulate the PDP scheme design as a throughput maximization problem as follows:

$$\begin{aligned} \max \quad & \sum_{k=1}^{K} \sum_{l \in \Re^k} \Upsilon_{k,l} \\ \text{s.t.:} \quad & (3) \sim \ (9),\ (11),\ (26),\ (27) \end{aligned}$$

where $\Upsilon_{k,l}$, λ_q, $f_{ij}^k(b, \xi_l)$, and $\tilde{f}_{ij}^k(\xi_n, \xi_l)$ are decision variables. Given the MISs of G, the above optimization problem is a linear programming.

Given the small values of $|\Re^k|$'s, the constraints (3) $\sim$ (9) introduce $O(KN)$ constraints and $O(KN^2)$ decision variables, where K is the number of CRVs and N is the number of CRSUs. The constraints (11), (26), and (27) introduce extra $O(N^2 + KN)$ constraints and extra $O(\mathcal{V})$ decision variables, where $\mathcal{V}$ is the number of MISs utilized in the linear programming. In this paper, these MISs are found through the scheduling index ordering (SIO)-based approach presented in [15], which implies that $\mathcal{V} = O(M^2(N^2 + KN)^2)$. In view of this, the formulated optimization is a linear programming with $O(N^2 + KN)$ constraints and $O(M^2(N^2 + KN)^2)$ decision variables, and the results in [20] implies that this optimization problem can be solved in polynomial time. Compared to the traditional scheme where multiple copies of the data are prefetched, the PDP scheme needs to determine how to split the prefetched data among different r-CRSUs. From (6)–(9), this operation introduces extra flow routing related constraints and variables ($O(KN)$ constraints and $O(KN^2)$ variables) to the throughput optimization problem formulated in this section. Based on previous discussions, the complexity of solving this throughput optimization problem is dominated by the link scheduling related constraints and variables. Thus, the extra operations introduced by the PDP scheme will not incur significant increase in computational complexity.

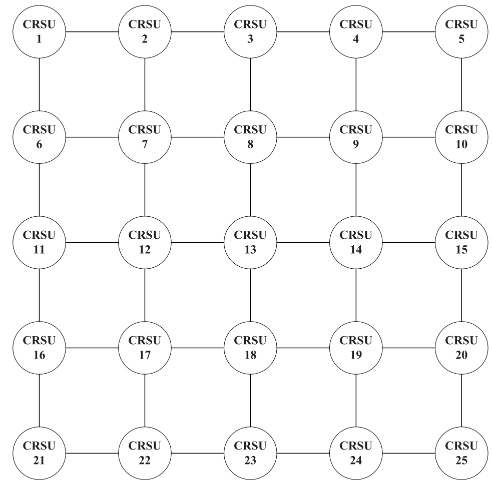

Fig. 5. Topology of the considered cell.

V. Performance Evaluation

A. Network Setup

We consider a cell with 25 CRSUs, which covers a 800 m $\times$ 800 m rectangular street area with a grid road layout. These CRSUs are deployed at intersections and the topology of the considered cell is shown in Fig. 5 where each pair of CRSUs is 200-m away. Among these CRSUs, CRSU13 is connected to data networks via wired connections and other CRSUs are wireless routers deployed to help CRVs get connections to data networks. CRSUs transmit at a power of 2 W and the transmission of a CRSU will cause nonnegligible interference to the reception of another CRSU when the received interference power exceeds 1.34×10^{-7} W. The wireless transmission is characterized by a path loss exponent of 3 and an antenna related constant of 4.63. The transmission between two CRSUs is successful when the power of the received signal is higher than 10^{-6} W. Thus, each CRSU has a transmission range of $R_T = 210$ m and an interference range of $R_I = 410$ m. Considering PUs' activities, the equivalent available bandwidth of band m, W_{ij}^m, is assumed to be a random variable uniformly distributed in $[0, 10]$ (MHz) [21]. In the following, α is set to 0.8 and $(T_k/T) = 0.6\ \forall k = 1, \ldots, K$. The maximum achievable data rate over 1 Hz bandwidth of the CR links is r_{CR} and that of the CV links is r_{CV}. For simplicity, we assume all CRSUs have the same number of radios H_r and all CRVs have the same number of radios H_v. At each intersection, a CRV goes straight, turns left, and turns right with probabilities p_0, p_1, and p_2. Note that these probabilities reflect how accurate the trajectories of CRVs can be predicted.

B. Results and Analysis

As discussed in Section VI, existing studies rely on reliable wired connections to ensure data delivery in case of inaccurate mobility prediction and thus is not applicable to the scenario where only unreliable wireless connections exist between r-CRSUs. Thus, in Fig. 6, the performance of the proposed PDP scheme is compared with that of the traditional approach where the SSP prefetches a copy of data intended

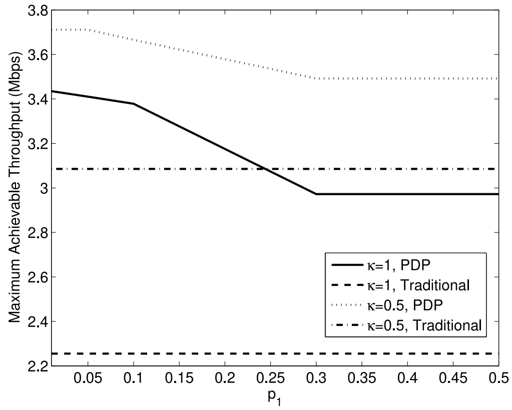

Fig. 6. Performance gain of the PDP scheme over the traditional approach.

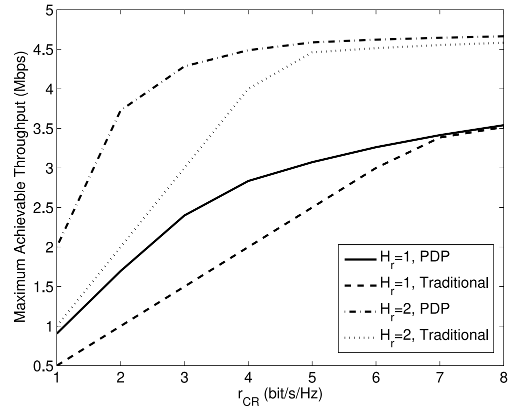

Fig. 7. Performance gain of the PDP scheme over the traditional approach with growing amount of available resources in the SWB.

for the kth CRV to each r-CRSU in $\Re^k$. We consider $K = 4$ CRVs and $M = 3$ harvested bands. Each CRSU is assumed to be equipped with two radios and each CRV has only one radio. κ in Fig. 6 represents the percentage of CRVs with $|\Re^k| > 1$. In other words, κ represents the percentage of CRVs whose mobility cannot be accurately predicted by the SSP. When $\kappa = 1$, we set $p_0 = 0$ and the 4 CRVs will potentially meet $\{CRSU1, CRSU6\}$, $\{CRSU21, CRSU22\}$, $\{CRSU25, CRSU20\}$, and $\{CRSU5, CRSU4\}$, respectively. The probabilities that those CRVs meet $CRSU1$, $CRSU21$, $CRSU25$, and $CRSU5$ are p_1, i.e., $p_1^1 = p_{21}^2 = p_{25}^3 = p_5^4 = p_1$. When $\kappa = 0.5$, the SSP is assumed to know two of those CRVs exactly meet $CRSU25$ and $CRSU5$, respectively, and all other settings remain the same with the $\kappa = 1$ case. r_{CR} is set to 3 bit/s/Hz and r_{CV} is set to 2 bit/s/Hz. From Fig. 6, the PDP scheme can achieve considerable performance gain over the traditional approach, which demonstrates the effectiveness of the PDP scheme. Note that the PDP scheme allows the SSP to take advantage of predicted contact probabilities to strategically reduce unnecessary downloading and adjustment flows. The more accurately the SSP can predict which CRSUs CRVs meet, the more network resources can be saved for data prefetching. This explains why the maximum achievable throughput of the PDP scheme slightly decreases with p_1 increasing in Fig. 6. It can be observed from Fig. 6 that, when $\kappa = 0.5$, the throughput gain of our PDP scheme could be less than 20%. This result is not surprising since, in these cases, the mobility of a large portion of CRVs can be accurately predicted. The advantage of our PDP scheme over the traditional approach is that our PDP scheme proactively exploits predicted contact probabilities to reduce unnecessary data prefetching due to the uncertain mobility of CRVs. Namely, the PDP scheme is designed to deal with the uncertainties in the mobility of CRVs and facilitate efficient data exchange in the V-CCHN. Its performance gain over the traditional approach decreases when the mobility of more CRVs can be accurately predicted, which explains why the performance gain is sometimes less than 20%. Moreover, as shown in Fig. 6, the maximum achievable throughput of the V-CCHN is much less than the rate expected to be supported by each link. This is resulted from interference coordination. Namely, the nodes in the considered network cannot always be active and thus the throughput of the whole network is much less than the average throughput of each node. This can be improved by equipping each nodes with multiple antennas and employing multiantenna techniques for interference cancellation, which allows nodes to be simultaneously active [22].

The effectiveness of the PDP scheme is further demonstrated in Fig. 7 where we examine how the performance of the PDP scheme and that of the traditional approach vary with the growing amount of network resources in the SWB, which are represented by r_{CR}. The parameters are the same as those in Fig. 6 other than $p_1 = 0.1$ and $M = 5$. In Fig. 7, the maximum achievable throughput is not only determined by the available resources in the SWB but also limited by the amount of resources available to CV links. Since the resources available to CV links remain unchanged when r_{CR} increases, the maximum achievable throughput approaches to an upper limit as r_{CR} growing, which explains the results in Fig. 7. Clearly from Fig. 7, with r_{CR} increasing, the maximum achievable throughput approaches this limit much faster under the PDP scheme. This implies that the extra resources added to the SWB can be more efficiently utilized under the PDP scheme, which justifies the design goal of PDP scheme. In addition, Fig. 7 shows that a higher maximum achievable throughput can be obtained when CRSUs are equipped with more radios. This result is very intuitive since CRSUs are able to exploit extra harvested spectrum bands with more radios.

How the maximum achievable throughput is related to r_{CR} and the amount of available spectrum resources is shown in Fig. 8. The parameters are the same as those in Fig. 6. Similar to the results in Fig. 7, the maximum achievable throughput increases with r_{CR} increasing and approaches an upper limit due to the constraints of CV links. Noticing that the SSP can exploit more harvested bands to facilitate data transmissions in the SWB and the CV links when $M = 5$, with r_{CR} increasing, we expect to observe both a larger initial growth rate and a higher limit of the maximum achievable throughput when $M = 5$, which is verified by the results in Fig. 8.

In Fig. 9, we study the relationship between the maximum achievable throughput and the amount of resources available to the CV links. The parameters are the same as

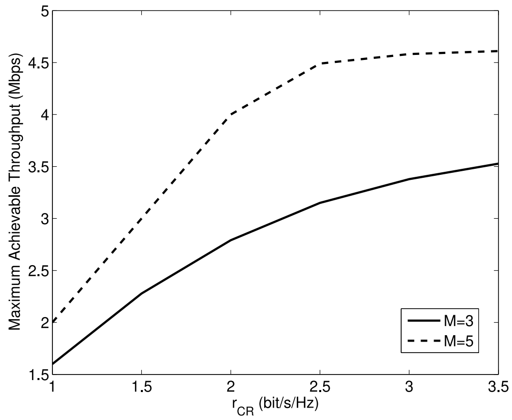

Fig. 8. Maximum achievable throughput versus r_{CR} and the amount of available spectrum resources.

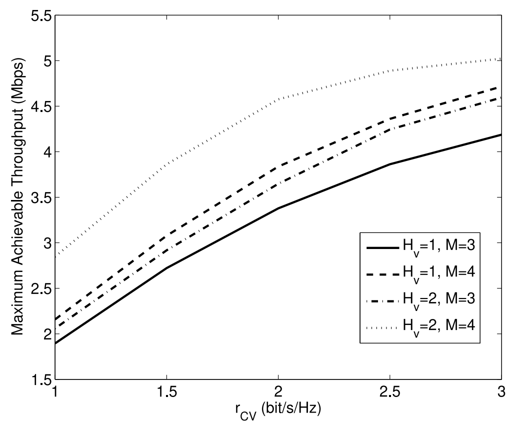

Fig. 9. Maximum achievable throughput versus the amount of resources available to the CV links.

in Fig. 6. It can be observed from Fig. 9 that the maximum achievable throughput increases with r_{CV} at a decreasing growth rate. Notice that the maximum achievable throughput depends on not only the data rates supported by the CV links but also how much data could be delivered via the SWB. The increases in the maximum achievable throughput indicate that it is initially limited by the data rate supported over the CV links. With r_{CV} increasing, the data rate supported on the CV links gradually improves and the data transmissions in the SWB become the bottleneck, which results in the decreasing growth rate shown in Fig. 9. According to Fig. 9, the maximum achievable throughput improves if either CRVs are equipped with more radios or more harvested spectrum bands are available to the SSP. This is not surprising since more resources are available to the SSP in either case.

VI. Related Work

Our V-CCHN architecture is first introduced in [5]. At the first glance, our proposal looks similar to traditional vehicular ad hoc networks (VANETs). However, our proposed V-CCHN is substantially different from traditional VANETs [23]–[27]. Traditional VANETs primarily study how to support safety and traffic related services through V2V and V2I communications within specific bands, such as the DSRC bands. Although existing works on VANETs have looked into Internet connectivity from vehicles, their data rate could not handle the traffic considered for V-CCHN since they typically address the problem based on their own licensed bands [28], [29]. In the V-CCHN, we proactively exploit the mobility of traveling vehicles as a medium, an opportunistic data carrier, for data transmissions, which offers us an alternative way to transport data from the location where it is collected to the place where it is consumed or utilized in a smart city environment, other than using wireless spectrum. Even though exiting works, such as [30], have used the mobility of vehicles to address the disconnection in VANETs, their design might not be efficient to support the services envisioned for the V-CCHN. The major challenge is that, due to the mobility of vehicles, the contact duration between vehicles and roadside infrastructures is usually short. The limited contact duration significantly limits the amount of data to be exchanged between vehicles and roadside infrastructures [31]. This is the reason why we take advantage of CR technologies to acquire extra bandwidth for short-range high-speed V2V and V2I communications in our design to make the V-CCHN effective. Unlike DSRC bands, the availability of the spectrum bands harvested through CR technologies is affected by the activities of PUs. Thus, the design intended for VANETs might not be efficient for the V-CCHN. A few works on VANETs assume that some nodes, such as vehicles and roadside infrastructures, are equipped with CRs [26], [27], [32]. Since these works primarily serve as the proof-of-concept study to evaluate the effectiveness of CRs in vehicular communications, newly designed schemes are still necessary.

Data prefetching has been combined with mobility prediction in various works to further improve the performance of data delivery in vehicular communications [9]–[12], [33]–[35]. Siris *et al.* [9] presented a distributed data prefetching approach which exploits user mobility information and a congestion pricing scheme. Zhang *et al.* [10] proposed the MobilityFirst future Internet architecture and the EdgeBuffer framework to enable efficient mobile content delivery. To achieve the design goal, they further develop a new mobility prediction model and a content prefetching protocol under the proposed architecture. Abani *et al.* [11] leveraged the ubiquity of caching nodes in information centric networks to address inaccurate mobility prediction and facilitate efficient data prefetching. Kim *et al.* [12] proposed vehicle route-based prefetching scheme to maximize data dissemination success probability by jointly considering the limited storage at roadside APs and the variations in wireless connectivity. Unlike these works where the infrastructure nodes are assumed to be reliably connected to data networks via wired connections, r-CRSUs in the V-CCHN are connected to data networks via the SWBs using harvested bands where spectrum availability is subject to PUs' activities. As a result, data prefetching in the V-CCHN is more challenging since the amount of data which can be delivered to vehicles not only limited by vehicle' mobility but also affected by the availability of harvested spectrum resources. That is why we develop the

PDP scheme for the V-CCHN so that unreliable harvested spectrum resources can be efficiently utilized to push data from infrastructures to CRVs.

VII. Conclusion

To fulfill the design goal of our V-CCHN, in this paper, we propose the PDP scheme so that the SSP can efficiently deliver the intended data to CRVs through the partial deployed infrastructure, CRSUs, and harvested spectrum resources for transmissions. We formulate the PDP schematic design as a throughput maximization problem where CVaR is used to handle the uncertainty in PUs' activities and CRVs' mobility. Through performance evaluation, we demonstrate that, when compared with the traditional data prefetching approach, our PDP scheme can achieve more efficient resource utilization within the SWB. The results also indicate that the performance of the V-CCHN improves with more available networking resources, such as harvested bands and radios per CRSU/CRV. We hope that this work will foster more research activities to exploit the unprecedent opportunities offered by the V-CCHN.

References

[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[2] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021*, Cisco, San Jose, CA, USA, Jun. 2017.

[3] N. Chen, M. Wang, N. Zhang, and X. Shen, "Energy and information management of electric vehicular network: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 967–997, 2nd Quart., 2020.

[4] H. Ding, Y. Fang, X. Huang, M. Pan, P. Li, and S. Glisic, "Cognitive capacity harvesting networks: Architectural evolution towards future cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1902–1923, 3rd Quart., 2017.

[5] H. Ding, C. Zhang, Y. Cai, and Y. Fang, "Smart cities on wheels: A newly emerging vehicular cognitive capability harvesting network for data transportation," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 160–169, Apr. 2018.

[6] H. Ding, X. Li, Y. Cai, B. Lorenzo, and Y. Fang, "Intelligent data transportation in smart cities: A spectrum-aware approach," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2598–2611, Dec. 2018.

[7] H. Ding and Y. Fang, "Virtual infrastructure at traffic lights: Vehicular temporary storage assisted data transportation at signalized intersections," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12452–12456, Dec. 2018.

[8] J. Feng, S. Yang, and Z. Feng, "Vehicle-assisted offloading on metropolitan streets: Enhancing geographical fluidity of wireless resources," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 622–625, Oct. 2017.

[9] V. A. Siris, X. Vasilakos, and G. C. Polyzos, "Efficient proactive caching for supporting seamless mobility," in *Proc. IEEE Int. Symp. World Wireless Mobile Multimedia Netw.*, 2014, pp. 1–6.

[10] F. Zhang *et al.*, "EdgeBuffer: Caching and prefetching content at the edge in the mobilityfirst future Internet architecture," in *Proc. IEEE 16th Int. Symp. World Wireless Mobile Multimedia Netw. (WoWMoM)*, 2015, pp. 1–9.

[11] N. Abani, T. Braun, and M. Gerla, "Proactive caching with mobility prediction under uncertainty in information-centric networks," in *Proc. 4th ACM Conf. Inf. Centric Netw.*, 2017, pp. 88–97.

[12] R. Kim, H. Lim, and B. Krishnamachari, "Prefetching-based data dissemination in vehicular cloud systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 292–306, Jan. 2016.

[13] Y. Ni *et al.*, "Toward reliable and scalable Internet of Vehicles: Performance analysis and resource management," *Proc. IEEE*, vol. 108, no. 2, pp. 324–340, Feb. 2020.

[14] M. Pan, P. Li, Y. Song, Y. Fang, P. Lin, and S. Glisic, "When spectrum meets clouds: Optimal session based spectrum trading under spectrum uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 615–627, Mar. 2014.

[15] H. Li, Y. Cheng, C. Zhou, and P. Wan, "Multi-dimensional conflict graph based computing for optimal capacity in MR-MC wireless networks," in *Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2010, pp. 774–783.

[16] H. Zhai and Y. Fang, "Impact of routing metrics on path capacity in multirate and multihop wireless ad hoc networks," in *Proc. IEEE Int. Conf. Netw. Protocols*, 2006, pp. 86–95.

[17] H. Yue, M. Pan, Y. Fang, and S. Glisic, "Spectrum and energy efficient relay station placement in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 883–893, May 2013.

[18] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *J. Bank. Finac.*, vol. 26, no. 7, pp. 1443–1471, Jul. 2002.

[19] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, no. 3, pp. 21–42, 2000.

[20] Y. Ye, "An $o(n^3 l)$ potential reduction algorithm for linear programming," *Math. Program.*, vol. 50, nos. 1–3, pp. 239–258, 1991.

[21] S. Yin, D. Chen, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: A large-scale spectrum measurement study," *IEEE Trans. Mobile Comput.*, vol. 11, no. 6, pp. 1033–1046, Jun. 2012.

[22] H. Ding, H. Yue, J. Liu, P. Si, and Y. Fang, "Energy-efficient secondary traffic scheduling with MIMO beamforming," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2015, pp. 1–6.

[23] J. Zhao and G. Cao, "VADD: Vehicle-assisted data delivery in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1910–1922, May 2008.

[24] J. He, L. Cai, J. Pan, and P. Cheng, "Delay analysis and routing for two-dimensional vanets using carry-and-forward mechanism," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1830–1841, Jul. 2017.

[25] H. He, H. Shan, A. Huang, and L. Sun, "Resource allocation for video streaming in heterogeneous cognitive vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7917–7930, Oct. 2016.

[26] D. Niyato, E. Hossain, and P. Wang, "Optimal channel access management with QoS support for cognitive vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 4, pp. 573–591, Apr. 2011.

[27] H. Zhou, N. Cheng, Q. Yu, X. S. Shen, D. Shan, and F. Bai, "Toward multi-radio vehicular data piping for dynamic dsrc/tvws spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2575–2588, Oct. 2016.

[28] T. Wang, L. Song, and Z. Han, "Coalitional graph games for popular content distribution in cognitive radio VANETs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 4010–4019, Oct. 2013.

[29] S. Yoon, D. T. Ha, H. Q. Ngo, and C. Qiao, "MOPADs: A mobility profile aided file downloading service in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5235–5246, Nov. 2009.

[30] K. C. Lee and M. Gerla, "Opportunistic vehicular routing," in *Proc. IEEE Eur. Wireless Conf. (EW)*, 2010, pp. 873–880.

[31] J. Zhao, X. Zhuo, Q. Li, W. Gao, and G. Cao, "Contact duration aware data replication in DTNs with licensed and unlicensed spectrum," *IEEE Trans. Mobile Comput.*, vol. 15, no. 4, pp. 803–816, Apr. 2016.

[32] M. Pan, P. Li, and Y. Fang, "Cooperative communication aware link scheduling for cognitive vehicular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 760–768, May 2012.

[33] N. Lu *et al.*, "Wi-Fi hotspot at signalized intersection: Cost-effectiveness for vehicular Internet access," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3506–3518, May 2016.

[34] W. Zhu, D. Li, and W. Saad, "Multiple vehicles collaborative data download protocol via network coding," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1607–1619, Apr. 2015.

[35] H. Zhou, N. Cheng, J. Wang, J. Chen, Q. Yu, and X. Shen, "Toward dynamic link utilization for efficient vehicular edge content distribution," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8301–8313, Sep. 2019.