This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Intelligent Data Transportation in Smart Cities: A Spectrum-Aware Approach

Haichuan Ding, Xuanheng Li, *Member, IEEE*, Ying Cai, Beatriz Lorenzo, *Member, IEEE*, and Yuguang Fang, *Fellow, IEEE*

***Abstract*— Communication technologies supply the blood for smart city applications. In view of the ever-increasing wireless traffic generated in smart cities and our already congested radio access networks (RANs), we have recently designed a *data transportation* network, the vehicular cognitive capability harvesting network (V-CCHN), which exploits the harvested *spectrum opportunity* and the *mobility opportunity* offered by the massive number of vehicles traveling in the city to not only offload delay-tolerant data from congested RANs but also support delay-tolerant data transportation for various smart-city applications. To make data transportation efficient, in this paper, we develop a *spectrum-aware* (SA) *data transportation* scheme based on Markov decision processes. Through extensive simulations, we demonstrate that, with the developed data transportation scheme, the V-CCHN is effective in offering data transportation services despite its dependence on dynamic resources, such as vehicles and harvested spectrum resources. The simulation results also demonstrate the superiority of the SA scheme over existing schemes. We expect the V-CCHN to well complement existing telecommunication networks in handling the exponentially increasing wireless data traffic.**

***Index Terms*— Smart cities, data transportation, data offloading, vehicular networks, cognitive radios.**

I. Introduction

THE initiatives on smart cities have offered us an informative and smart living environment where we can enjoy better and more convenient daily services, such as transportation, healthcare, and entertainment [1]. To support smart-city applications, numerous devices, such as sensors, cameras, and vehicles, are expected to connect and interact with each other for information sharing and delivery, intelligence extraction, and decision making, which will generate tremendous amount of wireless data traffic [2]–[5]. Although 4G/5G have great potential in dealing with huge mobile wireless data demands, they will face challenges in handling their promised services due to tremendous popularity of smart devices and soaring mobile applications (e.g., virtual reality and high definition video streaming) [6]. As a result, how to handle the huge amount of wireless data traffic is still challenging, particularly in a smart city environment [7], [8].

To address this challenge, we have recently designed a data transportation network, called vehicular cognitive capability harvesting network (V-CCHN), to support delay-tolerant data transportation for various smart-city applications [9]. Our basic idea is to employ vehicles traveling in cities as opportunistic data carriers to transport data from where it is collected to the places where it is consumed or utilized. Specifically, in the V-CCHN, data is delivered via the store-carry-forward mechanism by exploiting the opportunistic presence of vehicles and their mobility, under the supervision/management of a secondary service provider (SSP). Since the SSP might not gain full control of the mobility of these vehicles,[1] it often needs to count on a series of vehicles to carry and forward data in succession so that data can be delivered to intended locations. During this data delivery process, cognitive radio (CR) technologies are utilized to harvest spectrum resources for short-range high-speed data transmissions between vehicles. To facilitate efficient data delivery, the SSP collects various kinds of information, such as the availability of licensed/unlicensed bands, and makes spectrum allocation and data routing decisions to help data-carrying vehicles select data forwarding actions. The detailed introduction to the V-CCHN will be provided in Section III.A.

Noticing that the data transportation services of the V-CCHN is built on dynamic resources, such as harvested licensed/unlicensed bands, i.e., harvested bands, and the opportunistic presence of vehicles and their mobility, the SSP needs effective schemes to exploit these dynamic resources for data transportation.[2] Since only the conceptual

Manuscript received February 21, 2018; revised July 23, 2018 and September 10, 2018; accepted September 10, 2018; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor J. Huang. This work was supported by the U.S. National Science Foundation under Grant CNS-1717736 and Grant CNS-1409797. The work of X. Li was supported in part by the National Natural Science Foundation of China under Grant 61801080 and in part by the Dalian University of Technology through the Fundamental Research Funds under Grant DUT18RC(3)012. The work of Y. Cai was supported by the National Natural Science Foundation of China under Grant 61672106. The work of B. Lorenzo was supported in part by MINECO, Spain, under Grant TEC2016-76465-C2-2-R, in part by the Xunta de Galicia (Agrupación Estratéxica Consolidada de Galicia accreditation 2016–2019), and in part by the European Regional Development Fund. *(Corresponding author: Yuguang Fang.)*

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

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

Y. Cai is with the Computer School, Beijing Information Science and Technology University, Beijing 100101, China (e-mail: ycai@bistu.edu.cn).

B. Lorenzo is with the Department of Telematics, University of Vigo, 36310 Vigo, Spain (e-mail: blorenzo@gti.uvigo.es).

Digital Object Identifier 10.1109/TNET.2018.2871667

[1] In the V-CCHN, we only assume the SSP can supervise the operation of communication devices installed on vehicles. In the following, we will use vehicles and their installed communication devices interchangeably.

[2] Since this work is a preliminary study on the effectiveness of the V-CCHN, we assume enough vehicles participate, and leave the incentive-related issues for future works.

development of V-CCHN has been presented in [9], in this paper, we attempt to develop an effective data transportation scheme for the V-CCHN. Specifically, we seek for good data routing decisions at road intersections to fully exploit the harvested bands and the mobility of vehicles for data transportation. On the one hand, the data-carrying vehicles have more choices at intersections. On the other hand, the data routing decisions made at intersections determine the moving direction of carried data. If data routing decisions are not properly made, the SSP needs to dedicate extra resources to adjust data delivery and the corresponding data blocks could not even be delivered. With good data routing decisions at intersections, we could facilitate efficient operation of the V-CCHN and thus have a large-capacity data transportation network to complement existing telecommunications systems in handling the exponentially increasing wireless traffic generated from mobile and smart-city applications.

In this paper, under our V-CCHN, we carefully study how the SSP makes data routing decisions to help data-carrying vehicles select their data forwarding actions at intersections so that the considered data block can be efficiently delivered from the source to the destination. To make the data transportation processes effective, we introduce a *spectrum-aware (SA) data transportation* scheme where the effects of spectrum availability, the uncertain activities of licensed/unlicensed spectrum users, contention among different data-carrying vehicles, the mobility of the data-carrying vehicle, and the availability of relaying vehicles in each direction are jointly considered during the SSP's decision making process. We model the data delivery process as a Markov decision process (MDP) by observing that it involves a sequence of data routing decisions made at intersections. The optimal data routing decisions for the SSP are obtained via dynamic programming. Through extensive simulations, we thoroughly discuss the impacts of various parameters on the data delivery process. The results validate the effectiveness of our V-CCHN in handling the envisioned delay-tolerant data transportation services. Moreover, the results also demonstrate that, when compared with existing schemes, the SA approach can more efficiently support the data transportation in the V-CCHN.

II. Related Work

Data routing schematic design is an important research topic in vehicular ad hoc networks (VANETs) [10]–[13]. Noticing that VANETs are a special type of mobile ad hoc networks (MANETs), various MANET routing schemes, such as Greedy Perimeter Stateless Routing (GPSR), have been applied or extended to VANETs [14]. Unlike traditional MANETs, VANETs often lack continuous end-to-end (E2E) paths between the source and the destination [15]. To address this challenge, Zhu et al. propose two contact history based routing schemes by observing the temporal correlation of inter contact times (ICTs) between vehicles in [16] and [17]. A similar idea has been explored in [18] and [19] where utility-based routing schemes are designed by exploiting the patterns in the trajectories of vehicles. To efficiently utilize the constrained network capacity for data delivery, Wu *et al.* [20] develop a Capacity-Constrained Replication scheme where individual vehicles adjust their replication limits according to the varying network capacity. Clearly, all these schemes are developed with our experience of dealing with MANETs and delay tolerant networking (DTN). Different from traditional MANETs and DTN, data routing in VANETs is constrained by road layout, which allows us to develop more efficient routing schemes tailored to VANETs [14].

Darwish and Bakar [21] propose a Lightweight Intersection-based Traffic Aware Routing (LITAR) scheme for urban vehicular networks. In LITAR, using collector packets, vehicles measure the vehicular density and network connectivity on each road segment and make data routing decisions at intersections based on the collected information and the progress towards destinations. Zhang *et al.* [22] introduce the concept of link correlation and propose a routing metric called the expected transmission cost over a multi-hop path (ETCoP). Then, the authors develop a street-centric routing scheme where the data-carrying vehicles make routing decisions at intersections based on the estimated ETCoP of each road segment. Based on the concept of terminal intersections, Li *et al.* [23] design an adaptive quality-of-service (QoS)-based routing scheme for VANETs. With the proposed local QoS models of each segment, the optimal route is established through an ant colony optimization based algorithm. To address the local maximum problem in greedy-based routing, Togou *et al.* [24] propose to build a routing hierarchy consisting of stable backbones on road segment and bridge nodes at intersections. Based on this hierarchy, they further propose a stable CDS-based routing scheme where data routing decisions at each intersection are made based on the estimated data propagation delay on each road segment.

In [21]–[24], data routing decisions at intersections are primarily made based on the performance of data propagation along each road segment without considering the availability of vehicles at intersections. As a result, the schemes proposed therein might not be efficient due to the lack of vehicles at intersections [25]. This observation motivates a few recent works. Noticing that, at each intersection, only the road segments with vehicles on them can be used for packet forwarding, Darwish *et al.* [26] propose a reliable traffic aware routing scheme for VANETs where the routing decision at each intersection is made by jointly considering the traffic/network status and the availability of neighboring vehicles on each road segment. Similarly, He *et al.* [27] consider employing vehicles moving along the opposite direction to expedite data delivery. Noticing that the data might not be immediately directed along the desired directions due to the lack of vehicles at the considered intersection, they develop a model to estimate data transfer delays at the corresponding intersections and design a minimum delay routing algorithm (MDRA) accordingly.

When compared with existing works, spectrum uncertainty adds another dimension of randomness to the data routing in the V-CCHN, which further complicates the data routing schematic design. Different from traditional VANETs, the V-CCHN exploits CR technologies to harvest a wide range of under-utilized spectrum resources for short-range high-speed vehicle-to-vehicle transmissions. Unlike the spectrum

TABLE I
THE LIST OF IMPORTANT NOTATIONS AND DEFINITIONS

Notation	Definition
Δ	The size of the considered data block
K	The number of routing decisions can be made before the considered data block is discarded
$\overline{\Theta}$	A set of intersections which lead to the unsuccessful delivery of the considered data block
Θ	The complement of $\overline{\Theta}$
λ_n^v	The average number of potential relays along direction v at the nth intersection
p_n^v	The probability that a potential relay along direction v at the nth intersection is occupied by another data block
β_n^v	The number of potential relays along direction v at the nth intersection which are occupied by other data blocks
M_n	The total number of frequency bands at the nth intersection
ρ_n	The probability that a frequency band is available at the nth intersection
m_i	The number of bands requested to transfer the ith previously arrived data block
p_1, p_2, p_3,p_4	The probabilities that the data-carrying CRV keep current direction, turn left, turn right, and turn around at an interserction
$\mathcal{I}(k)$	The intersection where the data-carrying CRV locates at the kth decision epoch
$\vartheta(k)$	The moving direction of the data-carrying CRV at the kth decision epoch
$m(k)$	The number of harvested bands available to transfer the considered data block at the kth decision epoch
$\boldsymbol{\eta}(k)$	The availability of potential relays to receive the considered data block at the kth decision epoch
τ_j	The duration where the jth harvested band available for the considered data block can be used for secondary transmissions
λ_s	The mean of τ_j
T	The contact duration between two CRVs

licensed to vehicular communications, the availability of the spectrum resources harvested via CR technologies is subject to the activities of primary users (PUs). Due to the spatial and temporal variations in PUs' activities, the spectrum resources available at each intersection keep varying. Since the SSP might not be able to accurately predict PUs' future activities, it might not precisely know the future spectrum availability at subsequent intersections while making data routing decisions at the current intersection. Besides, it is hard for the SSP to exactly predict the spectrum available to each data block at subsequent intersections due to contention. To facilitate efficient data delivery in the V-CCHN, it is necessary for the SSP to consider this spectrum uncertainty at subsequent intersections when making data routing decisions at the current intersection. On the one hand, the performance of each data routing decision is closely related to data routing at subsequent intersections. On the other hand, because of the variations in spectrum availability, data might not be routed as expected at subsequent intersections due to the lack of enough spectrum resources and the mobility of the data-carrying vehicles at the corresponding intersections. Above discussions imply that, at each intersections, an efficient data delivery scheme for the V-CCHN needs to effectively handle not only the opportunistic presence of vehicles and their mobility but also the randomness incurred by the spectrum uncertainty at subsequent intersections. Since this spectrum uncertainty is generally not considered in existing works, the data routing schemes designed therein might not be efficient for data delivery in the V-CCHN.

III. SYSTEM MODEL

In this section, we provide an introduction to the V-CCHN architecture and elaborate on the routing problem as well as the

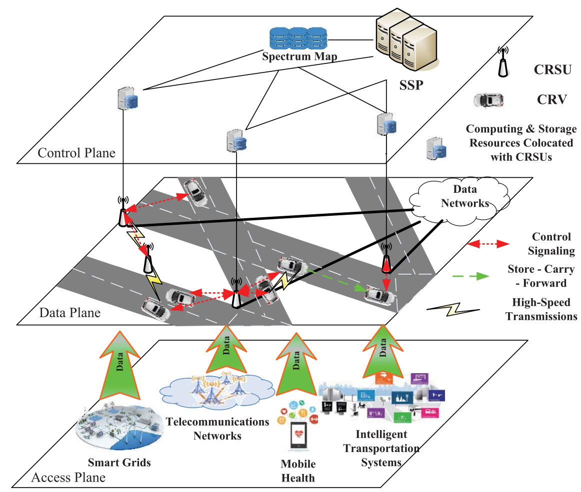

Fig. 1. The V-CCHN architecture.

corresponding models considered in this paper. The important notations are listed in Table I.

A. The V-CCHN Architecture

As shown in Fig. 1, the V-CCHN consists of an SSP, CR router enabled vehicles (CRVs), and CR capable roadside service units (CRSUs). The SSP is an independent wireless service provider with its own reliable bands (called basic bands in the subsequent development). For example, if cellular operators are the SSPs, the cellular bands can serve as the basic bands. The SSP recruits or deploys CRVs to provide delay-tolerant data transportation services in smart cities. CRSUs are the partial roadside infrastructures deployed by the SSP to improve the efficiency of data transportation. Generally speaking, there are two types of CRSUs in the V-CCHN. The first kind of CRSU does not have wired connections to data networks and are deployed by the SSP to deal with the uncertainty/dynamics in the V-CCHN and improve the efficiency of data transportation. For ease of presentation, this kind of CRSU will be called r-CRSU. The second kind of CRSU is called c-CRSU. They are deployed at strategic locations, such as vital intersections, and have wired connections to data networks, enabling the data exchange between the V-CCHN and data networks. These CRSUs can act as agents for the SSP to manage CRVs and r-CRSUs for data transportation in certain areas called cells.

The SSP's basic bands are primarily used to facilitate control signaling exchange between c-CRSUs and other network entities, such as CRVs and r-CRSUs, in corresponding cells for management and resource allocation. The management function of the SSP can be implemented in its deployed/leased fog nodes, c-CRSUs, or both. To make management more efficient, spectrum allocation can be implemented in c-CRSUs, whereas, the data routing decision making can be either implemented in the deployed/leased fog nodes or carried out by c-CRSUs. The requests for data transportation are first forwarded to, for example, a fog node in charge of a large geographic area. Then, the fog node determines whether to make routing decisions for these requests by itself or delegate the decision making tasks to a c-CRSU based on certain metrics, such as the distances to be traveled. After that, these routing decisions will be sent to data-carrying CRVs through c-CRSUs via the SSP's basic bands.

All CRVs and CRSUs are equipped with CR routers as communication devices. CR routers are powerful communication devices with agile communication interfaces, abundant computing resources and storage space. The agile communication interfaces of CR routers have cognitive radio (CR) capabilities and reconfigurability. Their CR capabilities allow CR routers to sense idle spectrum and exploit a wide range of under-utilized licensed/unlicensed spectrum for high-speed data transmissions. Their reconfigurability allows CR routers to exchange data with various types of end devices through the corresponding communication technologies, such as LTE, WiFi, and Bluetooth, used at these devices [28], [29]. With CR routers, CRVs can collect data from end devices when moving to their vicinities and collaboratively transport collected data to intended locations through the store-carry-forward mechanism and the harvested spectrum resources, for delivery or uploading.

B. Data Transportation in the V-CCHN

In the V-CCHN, the data transportation processes are supervised by the SSP.[3] Specifically, the SSP coordinates CRVs and CRSUs for spectrum sensing in order to build up spectrum map and collect spectrum statistics. With collected statistics, the SSP makes data routing decisions which help CRVs route

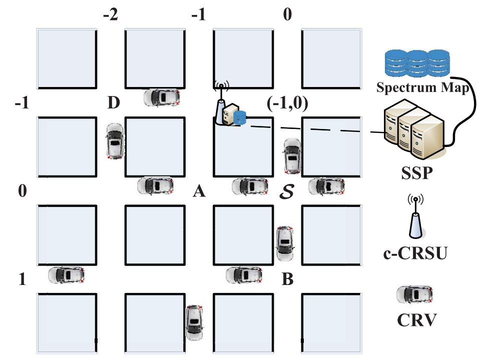

Fig. 2. An illustrative scenario. Intersection $\mathcal{S}$ locates at the 0th row and the 0th column. Intersection $(-1, 0)$ is the next intersection to the north of intersection $\mathcal{S}$.

data at various intersections with various spectrum and CRV availability and diverse levels of contentions. When selecting data forwarding actions at the intersections, the data-carrying CRVs query c-CRSUs in charge of the corresponding cells about available spectrum bands and determine whether to transfer data to another CRVs based on the data routing decisions received from the SSP. Once data has been transferred to another CRV, the next data-carrying CRV is responsible for carrying data towards the destination.[4] The aforementioned data transportation scheme is said to be spectrum-aware since data transmissions and spectrum availability are explicitly considered when the SSP makes data routing decisions.

C. Problem Setting

The generic data routing problem considered in this paper and its corresponding basic setting can be illustrated by the scenario shown in Fig. 2. A CRV approaches the intersection labeled $\mathcal{S}$ after collecting a data block of size Δ, such as a video clip from a surveillance camera, and the SSP makes data routing decisions to help CRVs select data forwarding actions at intersections so that this data block can be successfully delivered to the destination, such as a police station for the surveillance camera case, located around intersection D. The c-CRSU shown in Fig. 2 serves as the agent for the SSP to exchange control messages, such as data routing decisions, with CRVs. The roads considered here are either north-southward or east-westward. Thus, CRVs travel along directions $\upsilon \in \{N, S, W, E\}$, where N, S, W, and E represent north, south, west, and east, respectively.

In the V-CCHN, the SSP opportunistically exploits the presence of vehicles and their mobility for data transportation and thus might not be able to accurately know the trajectories of these vehicles other than their turning probabilities at intersections. In view of this, we assume that, at each intersection, a CRV keeps current direction, turns left, turns right, and turns around with probabilities p_1, p_2, p_3, and p_4, respectively [30]. Notice that the data block needs to travel through several intersections before arriving at the intersection D. To facilitate

[3]As aforementioned, the control functions of the SSP are not necessarily implemented in a single unit. We do not assume the existence of a single unit to control the whole network, but, we advocate the utilization of local coordinations for efficient resource allocation and data routing.

[4]Since this paper serves as a preliminary study on the data transportation in the V-CCHN, we assume that the data-carrying CRV discards the carried data once it has been successfully transferred to another CRV, for simplicity. How to exploit the redundancy offered by data duplication will be addressed elsewhere.

efficient data transportation, the SSP coordinates a series of CRVs to deliver this data block from intersection $\mathcal{S}$ to intersection D. At each intersection, the data-carrying CRV either keeps the data block or forwards it to another CRV traveling towards a certain direction away from the current intersection. For ease of presentation, the times when the routing decisions are made are called decision epochs. To facilitate efficient resource utilization and data delivery, the data block will be discarded if it is not delivered to intersection D after K decision epochs, with the decision made at intersection $\mathcal{S}$ as the first decision. On the basis of K and intersections $\mathcal{S}$ and D, we can divide all intersections into two groups, $\overline{\Theta}$ and Θ. If the considered data block is delivered to an intersection in $\overline{\Theta}$, it is impossible for it to be delivered to intersection D right after K decision epochs.[5] An intersection is in Θ if it is not in $\overline{\Theta}$. For example, if we set $K = 3$ for the scenario shown in Fig. 2, the intersection labeled A definitely belongs to Θ, whereas the intersection labeled B could belong to $\overline{\Theta}$. By definition, the data-carrying CRV can directly discard the data block without taking data forwarding actions if it reaches any intersection in $\overline{\Theta}$. Thus, we only need to study how the data-carrying CRV makes data routing decisions at intersections in Θ. For ease of presentation, the north-southward roads are called columns and the east-westward roads are called rows. Since only the north-southward and the east-westward roads are considered, each intersection can be represented by the indices of the row and column where it locates, taking intersection $\mathcal{S}$ located at the 0th row and the 0th column as a reference. For example, intersection $(-1, 0)$ is the next intersection to the north of intersection $\mathcal{S}$. Based on this representation, we can index the intersections in Θ as $\Theta_{\mathcal{I}} = \{1, \cdots, N\}$ following the lexicographic order, where N is the cardinality of Θ. The indices of intersection $\mathcal{S}$ and intersection D are denoted as n^s and n^d, where $n^s, n^d \in \Theta_{\mathcal{I}}$. In the following development, when referring to the nth intersection, we mean the nth intersection in $\Theta_{\mathcal{I}}$.

When approaching the nth ($n \in \Theta_{\mathcal{I}}$) intersection, the data-carrying CRV will query the c-CRSU about the CRVs available to relay the considered data block and the harvested bands available to transfer the data block to the relaying CRV. To ensure the contact duration is long enough to perform effective data transmissions, only those CRVs within a certain distance to the nth intersection could be considered as potential relaying CRVs. Denote the number of such CRVs moving towards $v \in \{N, S, W, E\}$ away from the nth intersection as α_n^v. These CRVs are called potential relays in the subsequent development. Similar to [25] and [27], α_n^v is assumed to be a Poisson random variable with mean λ_n^v. The total number of frequency bands at the nth intersection is M_n and each band is found to be available at a decision epoch with probability ρ_n. Noticing that the V-CCHN is designed to provide data transportation services to various smart-city applications, there could be multiple data blocks simultaneously handled at the nth intersection together with the considered one. When the data-carrying CRV requests resources for transferring the considered data block, the SSP could have already allocated certain number of potential relays and available bands to transfer previously arrived data blocks. In view of the random arrivals of data blocks at intersections, we assume that, when the data-carrying CRV arrives at the nth intersection, each potential relay moving towards direction v is occupied by another data block with probability p_n^v and m_i harvested bands are requested to transmit the ith of such data blocks. The number of such occupied potential relays along direction v is denoted as β_n^v. Then, the number of potential relays, along direction v, available to receive the considered data block is a Poisson random variable with mean $\lambda_n^v (1 - p_n^v)$ and β_n^v is a Poisson random variable with mean $\lambda_n^v p_n^v$. Considering the relatively short contact duration and large data sizes, each potential relay can be designated to relay at most one data block.

[5]How to find such a $\overline{\Theta}$ will be illustrated in the simulation study.

IV. The Optimal Data Routing Decisions

Based on above settings, we formulate the delivery process of the considered data block as an MDP, which allows the SSP to obtain the optimal data routing decisions via dynamic programming.

A. State Space and Actions

Since the routing of the considered data block depends on the specific intersection where the data-carrying CRV stays as well as the spectrum availability and the number of potential relays at the corresponding intersection, the state of the considered system at the kth decision epoch is

$$s(k) = (\mathcal{I}(k), \vartheta(k), m(k), \boldsymbol{\eta}(k)), \tag{1}$$

where $\mathcal{I}(k)$ is the intersection where the data-carrying CRV locates at the kth decision epoch, $\vartheta(k)$ is the current moving direction of the data-carrying CRV, $m(k)$ is the number of harvested bands available for the transferring of the considered data block, $\boldsymbol{\eta}(k) = \left[\eta_S(k)\ \eta_N(k)\ \eta_E(k)\ \eta_W(k)\right]^{\mathrm{T}}$, T signifies the transpose of a matrix, $\eta_v(k) = 1$ means at least a potential relay, traveling along the direction v, is available to receive the considered data block, and $\eta_v(k) = 0$, otherwise. By definition, the considered data block cannot be delivered to intersection D in K decision epochs once it is delivered to an intersection in $\overline{\Theta}$. Namely, all intersections in $\overline{\Theta}$ correspond to the case where the considered data block cannot be delivered. In view of this, we represent all intersections in $\overline{\Theta}$ as a virtual intersection and index it as the $(N+1)$th intersection. If the considered data block is delivered to an intersection in $\overline{\Theta}$ after a decision epoch, we say it is delivered to the $(N + 1)$th intersection, and vice versa. On the other hand, when the considered data block is delivered to the destination, it becomes delivered. In both cases, the data delivery process terminates. When reflecting in the formulation, $s(k)$'s with $\mathcal{I}(k) = N+1$ or $\mathcal{I}(k) = n^d$ will not transfer to any other state. In the subsequent development, we assume that the activities of licensed/unlicensed users on different bands share the same statistical properties for simplicity, which explains why $m(k)$ is used in $s(k)$. Our formulation can cover the case where the activities of licensed/unlicensed users on different bands have different statistical properties by introducing separate state variables to $s(k)$ for different bands. However, this would enlarge the state space. Since we aim to evaluate the effectiveness of the V-CCHN in this paper, we do not consider this diversity in frequency bands and leave it as a future work.

At the kth decision epoch, the SSP makes the data routing decision for the considered data block, denoted as $a\left(k\right)$, according to $s\left(k\right)$. In general, the SSP has the following choices for $a\left(k\right)$

$$a\left(k\right) = \begin{cases} 0 & \text{Keep the data block in the storage} \\ \delta & \text{Forward the data block to } \delta, \end{cases} \tag{2}$$

where $\delta = \{N, S, W, E\}$ represents a potential relay moving along the direction indicated by δ. The action δ is available to the SSP at the kth decision epoch only when $\eta_\delta\left(k\right) = 1$, $m\left(k\right) > 0$. The set of rules used by the SSP in selecting $a\left(k\right)$, $k = 1, \cdots, K$ is called a policy, denoted as g. By solving this MDP, the SSP can obtain the data routing policy, denoted as g^*, which is used to assist data-carrying CRVs in selecting data forwarding actions.

B. State Transition Probabilities

By definition, the probability that state $s\left(k\right)$ transfers to state $s\left(k+1\right)$ under data routing decision $a\left(k\right)$ can be formulated as shown in (3), as shown at the bottom of this page, where $\vartheta \in \{N, S, W, E\}$. $n' \in \{n_N, n_S, n_W, n_E\}$, n_N, n_S, n_W, and n_E are the indices of the intersections to the north, south, west, and east of the nth intersection, respectively. Since the number of potential relays available to receive the considered data block is independent of the spectrum availability at each intersection, the first term to the righthand side of (3), denoted as ϕ, can be reformulated as (4), as shown at the bottom of this page.

Notice that data routing decisions at different decision epochs could be made at different intersections. Considering the spatial variations in the activities of licensed/unlicensed users, we assume that the availability of a specific band at each epoch is independent.[6] Noticing that the data-carrying CRV might meet different sets of potential relays at different intersections, $\boldsymbol{\eta}\left(k\right)$'s are assumed to be independent for this study. Then, ϕ can be expressed as

$$\phi = \underbrace{\mathrm{P}\left(m\left(k+1\right) = m' \left| \mathcal{I}\left(k+1\right) = n' \right.\right)}_{\triangleq \phi_1} \times \underbrace{\mathrm{P}\left(\boldsymbol{\eta}\left(k+1\right) = \boldsymbol{\eta}' \left| \mathcal{I}\left(k+1\right) = n' \right.\right)}_{\triangleq \phi_2}. \tag{5}$$

The probabilities on the righthand side of (5) is conditioned on $\mathcal{I}\left(k+1\right)$ since the distributions of $m\left(k+1\right)$ and $\boldsymbol{\eta}\left(k+1\right)$ depend on which intersection the data-carrying CRV stays.

In next few subsections, ϕ and φ in (3) will be derived separately. Since the derivation of state transition probabilities when $\mathcal{I}\left(k\right) = n^d$ and $\mathcal{I}\left(k\right) = N+1$ is trivial, we will only derive φ for the case when $\mathcal{I}\left(k\right) \neq n^d$ and $\mathcal{I}\left(k\right) \neq N+1$ in the following analysis.

C. Derivation of ϕ

From (5), to obtain ϕ, we should derive ϕ_1 and ϕ_2. According to (5), ϕ_1 is the conditional distribution of $m\left(k+1\right)$ given $\mathcal{I}\left(k+1\right) = n'$. By definition, $m\left(k+1\right)$ equals the difference between the total number of harvested bands, $M\left(k+1\right)$, and the number of harvested bands allocated to transfer the scheduled data blocks, i.e.,

$$m\left(k+1\right) = \max\left\{M\left(k+1\right) - \sum_{v\in V}\sum_{i=1}^{\beta_{n'}^v} m_i, 0\right\}, \tag{6}$$

where $\beta_{n'}^v$ is the number of potential relays occupied by other data blocks, and m_i is the number of harvested bands requested to transmit the ith of these data blocks. The subscript of $\beta_{n'}^v$ is n' since we focus on the derivation of ϕ_1 which is the conditional distribution of $m\left(k+1\right)$ given $\mathcal{I}\left(k+1\right) = n'$. The max operation is adopted in (6) to include the case where no harvested band is available to the considered data block. Since different data blocks could have diverse data sizes, the number of harvested bands required to transfer these data blocks could be different. In view of this, we model m_i's as random variables. Notice that both $M\left(k+1\right)$ and the summation term in (6), $\theta = \sum\limits_{v\in V}\sum\limits_{i=1}^{\beta_{n'}^v} m_i$, are random variables. To obtain the conditional distribution of $m\left(k+1\right)$, we should derive the conditional distributions of $M\left(k+1\right)$ and θ.

[6]This independence assumption on spectrum availability can be relaxed, and only (7) in our formulation needs to be revised since, in this case, the number of harvested bands at the $(k+1)$th decision epoch will possibly also depend on the number of harvested bands at the kth decision epoch.

$$\begin{aligned} &\mathrm{P}\left(s\left(k+1\right) = \left(n', \vartheta', m', \boldsymbol{\eta}'\right) \left| s\left(k\right) = \left(n, \vartheta, m, \boldsymbol{\eta}\right), a(k) = a \right.\right) \\ &= \underbrace{\mathrm{P}(m\left(k+1\right) = m', \boldsymbol{\eta}\left(k+1\right) = \boldsymbol{\eta}' \left| \mathcal{I}\left(k\right) = n, \vartheta\left(k\right) = \vartheta, m\left(k\right) = m, \boldsymbol{\eta}\left(k\right) = \boldsymbol{\eta}, \mathcal{I}\left(k+1\right) = n', \vartheta\left(k+1\right) = \vartheta', a(k) = a \right.)}_{\triangleq \phi} \\ &\quad \times \underbrace{\mathrm{P}(\mathcal{I}\left(k+1\right) = n', \vartheta\left(k+1\right) = \vartheta' \mid \mathcal{I}\left(k\right) = n, \vartheta\left(k\right) = \vartheta, m\left(k\right) = m, \boldsymbol{\eta}\left(k\right) = \boldsymbol{\eta}, a(k) = a)}_{\triangleq \varphi}. \end{aligned} \tag{3}$$

$$\begin{aligned} \phi = &\ \mathrm{P}\left(m\left(k+1\right) = m' \left| \mathcal{I}\left(k\right) = n, \vartheta\left(k\right) = \vartheta, m\left(k\right) = m, \boldsymbol{\eta}\left(k\right) = \boldsymbol{\eta}, \mathcal{I}\left(k+1\right) = n', \vartheta\left(k+1\right) = \vartheta', a\left(k\right) = a \right.\right) \\ &\times \mathrm{P}\left(\boldsymbol{\eta}\left(k+1\right) = \boldsymbol{\eta}' \left| \mathcal{I}\left(k\right) = n, \vartheta\left(k\right) = \vartheta, m\left(k\right) = m, \boldsymbol{\eta}\left(k\right) = \boldsymbol{\eta}, \mathcal{I}\left(k+1\right) = n', \vartheta\left(k+1\right) = \vartheta', a\left(k\right) = a \right.\right). \end{aligned} \tag{4}$$

Given $\mathcal{I}(k+1) = n'$, $M(k+1)$ follows a binomial distribution with parameters $M_{n'}$ and $\rho_{n'}$, i.e.,

$$\begin{aligned} &\mathrm{P}\left(M(k+1) = M \left| \mathcal{I}(k+1) = n' \right.\right) \\ &\quad = \binom{M_{n'}}{M} \rho_{n'}^{M} (1-\rho_{n'})^{M_{n'}-M}, \quad 0 \le M \le M_{n'}. \end{aligned} \tag{7}$$

θ is the summation of m_i's. In practice, the SSP can learn the distributions of m_i's by gathering corresponding statistics. Since such distributions are currently not available and the number of bands required to handle a data block is closely related to its size, we assume that m_i follows a truncated Poisson distribution with parameter λ to facilitate the data routing policy design [31], i.e.,

$$\mathrm{P}(m_i = m) = \begin{cases} \frac{\lambda^m}{1-e^{-\lambda}} \dfrac{e^{-\lambda}}{m!} & m \ge 1 \\ 0 & m = 0. \end{cases} \tag{8}$$

Then, θ is the summation of $\sum\limits_{v \in V} \beta_{n'}^{v}$ random variables with the distribution shown in (8). Since $\beta_{n'}^{v}$'s are Poisson random variables, $\sum\limits_{v \in V} \beta_{n'}^{v}$ is a Poisson random variable with mean $\Lambda = \sum\limits_{v \in V} \lambda_{n'}^{v} p_{n'}^{v}$. Based on the law of total probability, the distribution of θ can be expressed as

$$\begin{aligned} &\mathrm{P}\left(\theta = \ell \left| \mathcal{I}(k+1) = n' \right.\right) \\ &\quad = \begin{cases} \sum\limits_{l=1}^{\infty} \dfrac{e^{-\Lambda}\Lambda^l}{l!} \mathrm{P}\left(\sum\limits_{i=1}^{l} m_i = \ell\right) & \ell \ge 1 \\ e^{-\Lambda} & \ell = 0. \end{cases} \end{aligned} \tag{9}$$

To facilitate the derivation of $\mathrm{P}\left(\theta = \ell \left| \mathcal{I}(k+1) = n' \right.\right)$, the distribution of $\iota = \sum\limits_{i=1}^{l} m_i$, the sum of l independently identically distributed (i.i.d.) random variables with distribution functions shown in (8), is obtained in the following lemma.

Lemma 1: The distribution of ι is

$$\mathrm{P}(\iota = \ell) = \frac{\lambda^\ell e^{-l\lambda}}{\ell!(1-e^{-\lambda})^l} \sum_{i=0}^{l} \binom{l}{i} (-1)^{l-i} i^\ell. \tag{10}$$

Proof: Noticing that $\mathrm{P}(\iota = \ell)$'s are probability mass functions and $0 \le \mathrm{P}(\iota = \ell) \le 1$, the series $\sum\limits_{\ell=-\infty}^{\infty} \mathrm{P}(\iota = \ell) z^\ell$ uniformly converges in the disc $|z| < 1$ on the complex plane, where $z \in \mathbb{C}$ is a complex number [32]. Then, with the probability mass function of ι, we can define an analytic function on the disc $\{z \in \mathbb{C} \,||z| < 1\}$ as

$$\begin{aligned} X_\iota(z) &= \sum_{\ell=-\infty}^{\infty} \mathrm{P}(\iota = \ell) z^\ell = \mathrm{E}\left[z^\iota\right] \\ &= \mathrm{E}\left[z^{\sum\limits_{i=1}^{l} m_i}\right] = \prod_{i=1}^{l} \mathrm{E}\left[z^{m_i}\right], |z| < 1. \end{aligned} \tag{11}$$

With (8), we have

$$\begin{aligned} \mathrm{E}\left[z^{m_i}\right] &= \frac{1}{1-e^{-\lambda}} \sum_{\ell=1}^{\infty} \frac{e^{-\lambda}\lambda^\ell}{\ell!} z^\ell \\ &= \frac{e^{-\lambda+\lambda z}}{1-e^{-\lambda}} \sum_{\ell=1}^{\infty} \frac{(\lambda z)^\ell}{\ell!} e^{-\lambda z} \\ &= \frac{e^{-\lambda}\left(e^{\lambda z}-1\right)}{1-e^{-\lambda}}. \end{aligned} \tag{12}$$

Plugging (12) into (11), it follows

$$\begin{aligned} X_\iota(z) &= \frac{e^{-l\lambda}\left(e^{\lambda z}-1\right)^l}{\left(1-e^{-\lambda}\right)^l} \\ &= \frac{e^{-l\lambda}}{\left(1-e^{-\lambda}\right)^l} \sum_{i=0}^{l} \binom{l}{i} (-1)^{l-i} e^{i\lambda z}, \quad |z| < 1. \end{aligned} \tag{13}$$

With Cauchy's Integral Formula for Derivatives, $\mathrm{P}(\iota = \ell)$ can be rewritten as [32]

$$\begin{aligned} \mathrm{P}(\iota = \ell) &= \frac{1}{2\pi \mathrm{i}} \oint_\Gamma \frac{X_\iota(z)}{z^{\ell+1}} dz \\ &= \frac{e^{-l\lambda}}{\left(1-e^{-\lambda}\right)^l} \sum_{i=0}^{l} \binom{l}{i} (-1)^{l-i} \frac{1}{2\pi \mathrm{i}} \oint_\Gamma \frac{e^{i\lambda z}}{z^{\ell+1}} dz, \end{aligned} \tag{14}$$

where $\mathrm{i} = \sqrt{-1}$, Γ is a circle centered at the origin followed the anticlockwise direction with radius less than 1. Clearly, $e^{i\lambda z}$ is analytic in the disc $\{z \in \mathbb{C} \,||z| < 1\}$. Thus, $e^{i\lambda z}/z^{\ell+1}$ has a pole of order $\ell + 1$ at 0. Following from the Residue Theorem, (14) can be reformulated as [32]

$$\mathrm{P}(\iota = \ell) = \frac{e^{-l\lambda}}{\left(1-e^{-\lambda}\right)^l} \sum_{i=0}^{l} \binom{l}{i} (-1)^{l-i} \mathrm{res}\left(\frac{e^{i\lambda z}}{z^{\ell+1}}, 0\right), \tag{15}$$

where $\mathrm{res}\left(e^{i\lambda z}/z^{\ell+1}, 0\right)$ is the residue of $e^{i\lambda z}/z^{\ell+1}$ at 0. By definition of the residue, we have

$$\mathrm{res}\left(\frac{e^{i\lambda z}}{z^{\ell+1}}, 0\right) = \frac{1}{\ell!} \lim_{z \to 0} \frac{d^\ell}{dz^\ell} e^{i\lambda z} = \frac{1}{\ell!} (i\lambda)^\ell. \tag{16}$$

With (15) and (16), $\mathrm{P}(\iota = \ell)$ can be derived as shown in (10). ■

Remark: According to 0.154 in [33], we have

$$\begin{aligned} \sum_{i=0}^{l} \binom{l}{i} (-1)^{l-i} i^\ell &= (-1)^l \sum_{i=0}^{l} \binom{l}{i} (-1)^i i^\ell \\ &= 0, \quad \forall l \ge \ell + 1. \end{aligned} \tag{17}$$

This result matches our intuition that $\ell \ge l$, i.e., the simultaneous transmissions of l data blocks require at least l harvested bands. With (17), $\mathrm{P}(\iota = \ell)$ can be reformulated as

$$\begin{aligned} &\mathrm{P}(\iota = \ell) \\ &\quad = \begin{cases} \frac{\lambda^\ell e^{-l\lambda}}{\ell!(1-e^{-\lambda})^l} \sum\limits_{i=0}^{l} \binom{l}{i} (-1)^{l-i} i^\ell & \ell \ge l \\ 0 & \mathit{Otherwise}. \end{cases} \end{aligned} \tag{18}$$

With Lemma 1, $\mathrm{P}\left(\theta=\ell\left|\mathcal{I}\left(k+1\right)=n'\right.\right)$ can be derived in the following theorem.

Theorem 1: The distribution of θ is

$$\mathrm{P}\left(\theta=\ell\left|\mathcal{I}\left(k+1\right)=n'\right.\right) = \begin{cases} \frac{e^{-\Lambda}\lambda^{\ell}}{\ell!}\sum\limits_{l=1}^{\ell}\sum\limits_{i=0}^{l}\frac{(-1)^{l-i}\Lambda^{l}i^{\ell}}{(l-i)!i!(e^{\lambda}-1)^{l}} & \ell\geq 1 \\ e^{-\Lambda} & \ell=0. \end{cases} \tag{19}$$

Proof: The distribution of θ directly follows from (9) and (18). ∎

From (6), the conditional distribution of $m\left(k+1\right)$ given $\mathcal{I}(k+1)=n'$ can be expressed as shown in (20), at the bottom of this page. With (19), we have (21), as shown at the bottom of this page. When $\{m=0,M=0\}$, $\{m=M,M\geq 1\}$, and $\{m=0,M=1\}$, the conditional probability in (21) equals 1, $e^{-\Lambda}$, and $1-e^{-\Lambda}$, respectively.

Plugging (7) and (21) into (20), ϕ_1 can be finally obtained.

By definition, ϕ_2 is the conditional distribution of $\boldsymbol{\eta}'=\left[\,\eta_N\ \eta_S\ \eta_W\ \eta_E\,\right]^{\mathrm{T}}$ given $\mathcal{I}(k+1)=n'$ and thus can be derived as

$$\phi_2=\prod_{\upsilon\in\{N,S,W,E\}}\Big\{\left(1-\eta_{\upsilon}'\right)e^{-\lambda_{n'}^{\upsilon}\left(1-p_{n'}^{\upsilon}\right)} +\eta_{\upsilon}'\left(1-e^{-\lambda_{n'}^{\upsilon}\left(1-p_{n'}^{\upsilon}\right)}\right)\Big\}. \tag{22}$$

Plugging the expressions for ϕ_1 and ϕ_2 back to (5), ϕ can finally be derived.

D. Derivation of φ

Considering the constraint on road layout, if the data-carrying CRV moves towards the direction of ϑ' at the $(k+1)$th decision epoch, the n'th intersection must lie to the direction of ϑ' of the nth intersection. Namely, given ϑ', it must have $n'=n_{\vartheta'}$, which implies that $\varphi=0$ if $n'\neq n_{\vartheta'}$. Thus, in the following, we only focus on the case where $n'=n_{\vartheta'}$. In this case, φ can be reformulated as

$$\begin{aligned}\varphi=&\mathrm{P}(\mathcal{I}\left(k+1\right)=n_{\vartheta'},\vartheta\left(k+1\right)=\vartheta'\left|\mathcal{I}\left(k\right)=n\right.,\\ &\vartheta\left(k\right)=\vartheta,m\left(k\right)=m,\boldsymbol{\eta}\left(k\right)=\boldsymbol{\eta},a(k)=a).\end{aligned} \tag{23}$$

To obtain the expression of φ, we need to consider the following cases.

1) $a=0$: In this case, the data-carrying CRV keeps the data block in its own storage and thus the data block moves to the intended intersection only when the data-carrying CRV turns to the corresponding direction. In view of this, φ, for the considered case, can be derived as shown in (24), at the bottom of this page.

2) $a=\delta,\vartheta'\neq\delta$: When $a(k)=\delta$, the data-carrying CRV tries to deliver data to a potential relay traveling along the direction δ. In this case, $\vartheta'\neq\delta$ only when the transmission attempt fails, and ϑ' results from the turn of the data-carrying CRV. Thus, φ is the probability that the data transmission fails and the data-carrying CRV turns to the direction ϑ'. Let $\mathcal{F}$ be the event that the transmission of the considered data block fails, given $\mathcal{I}\left(k\right)=n$, $\vartheta\left(k\right)=\vartheta$, $m\left(k\right)=m$, $\eta\left(k\right)=\eta$, $a\left(k\right)=\delta$, and $\mathcal{T}_{\vartheta'}$ be the event that the data-carrying CRV turns to the direction ϑ', give $\mathcal{I}\left(k\right)=n$, $\vartheta\left(k\right)=\vartheta$, $m\left(k\right)=m$, $\eta\left(k\right)=\eta$, $a\left(k\right)=\delta$. Then, the event $\mathcal{I}\left(k+1\right)=n_{\vartheta'},\vartheta\left(k+1\right)=\vartheta'$ given $\mathcal{I}\left(k\right)=n$, $\vartheta\left(k\right)=\vartheta$, $m\left(k\right)=m$, $\eta\left(k\right)=\eta$, $a\left(k\right)=\delta$ is equivalent to $\mathcal{F}\cap\mathcal{T}_{\vartheta'}$. In view of this, we can rewrite φ through $\mathcal{F}$ and $\mathcal{T}_{\vartheta'}$ as

$$\varphi=\mathrm{P}\left(\mathcal{F}\cap\mathcal{T}_{\vartheta'}\right). \tag{25}$$

Noticing that the moving direction of the data-carrying CRV is independent of the transmission of the considered data block, it follows

$$\varphi=\mathrm{P}\left(\mathcal{F}\right)\mathrm{P}\left(\mathcal{T}_{\vartheta'}\right). \tag{26}$$

As aforementioned, $a(k)=\delta$ implies that $\eta_{\delta}\left(k\right)=1$, $m\left(k\right)>0$. In this case, the transmission failure is caused

---

$$\mathrm{P}\left(m\left(k+1\right)=m\left|\mathcal{I}\left(k+1\right)=n'\right.\right)=\sum_{M=0}^{M_{n'}}\Big\{\mathrm{P}\left(\max\left\{M\left(k+1\right)-\theta,0\right\}=m\left|M\left(k+1\right)=M,\mathcal{I}\left(k+1\right)=n'\right.\right) \times\mathrm{P}\left(M\left(k+1\right)=M\left|\mathcal{I}\left(k+1\right)=n'\right.\right)\Big\}. \tag{20}$$

$$\begin{aligned}&\mathrm{P}\left(\max\left\{M\left(k+1\right)-\theta,0\right\}=m\left|M\left(k+1\right)=M,\mathcal{I}\left(k+1\right)=n'\right.\right)\\ &=\begin{cases}\mathrm{P}\left(\theta=M-m\left|\mathcal{I}\left(k+1\right)=n'\right.\right) & 1\leq m<M,M\geq 1\\ \left(1-\sum\limits_{\ell=0}^{M-1}\mathrm{P}\left(\theta=\ell\left|\mathcal{I}\left(k+1\right)=n'\right.\right)\right) & m=0,M>1\end{cases}\\ &=\begin{cases}\dfrac{e^{-\Lambda}\lambda^{(M-m)}}{(M-m)!}\sum\limits_{l=1}^{M-m}\sum\limits_{i=0}^{l}\dfrac{(-1)^{l-i}\Lambda^{l}i^{M-m}}{(l-i)!i!(e^{\lambda}-1)^{l}} & 1\leq m<M,M\geq 1\\ 1-e^{-\Lambda}-\sum\limits_{\ell=1}^{M-1}\dfrac{e^{-\Lambda}\lambda^{\ell}}{\ell!}\sum\limits_{l=1}^{\ell}\sum\limits_{i=0}^{l}\dfrac{(-1)^{l-i}\Lambda^{l}i^{\ell}}{(l-i)!i!(e^{\lambda}-1)^{l}} & m=0,M>1.\end{cases}\end{aligned} \tag{21}$$

$$\varphi=\begin{cases}p_1 & \{\vartheta=N,\vartheta'=N\}\cup\{\vartheta=S,\vartheta'=S\}\cup\{\vartheta=W,\vartheta'=W\}\cup\{\vartheta=E,\vartheta'=E\}\\ p_2 & \{\vartheta=N,\vartheta'=W\}\cup\{\vartheta=S,\vartheta'=E\}\cup\{\vartheta=W,\vartheta'=S\}\cup\{\vartheta=E,\vartheta'=N\}\\ p_3 & \{\vartheta=N,\vartheta'=E\}\cup\{\vartheta=S,\vartheta'=W\}\cup\{\vartheta=W,\vartheta'=N\}\cup\{\vartheta=E,\vartheta'=S\}\\ p_4 & \{\vartheta=N,\vartheta'=S\}\cup\{\vartheta=S,\vartheta'=N\}\cup\{\vartheta=W,\vartheta'=E\}\cup\{\vartheta=E,\vartheta'=W\}\end{cases} \tag{24}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

by the lack of enough available spectrum resources to support the delivery of the considered data block. Without loss of generality, we assume that each harvested band has a bandwidth of w and transmissions on these harvested bands can achieve a spectral efficiency of c. Due to the uncertain activities of lincensed/unlicensed users, the duration where each harvested band can be utilized for transmissions subject to variations and thus is modeled by a random variable τ_j, where the subscript j represents the jth harvested band available to transfer the considered data block. Then, $\mathrm{P}(\mathcal{F})$ can be expressed as

$$\mathrm{P}(\mathcal{F}) = \mathrm{P}\left(\sum_{j=1}^{m(k)} cw \times \min\{\tau_j, T\} < \Delta\right), \tag{27}$$

where $m(k)$ is the number of harvested bands available for transferring the considered data block of size Δ, T is the contact duration between two CRVs. The min operation in (27) implies that the data transmissions between two CRVs cannot last longer than the contact duration. Similar to [34], τ_j's are assumed to be i.i.d. exponential random variables with parameter λ_s. Then, $\mathrm{P}(\mathcal{F})$ can be derived as shown in the following Theorem.

Theorem 2: When $\frac{\Delta}{cw} > m(k)T$, $\mathrm{P}(\mathcal{F}) = 1$. When $\frac{\Delta}{cw} \le m(k)T$, $\mathrm{P}(\mathcal{F})$ can be derived in a closed-form as

$$\begin{aligned}\mathrm{P}(\mathcal{F}) = &\sum_{j=0}^{\lfloor \frac{\Delta}{cw}/T \rfloor} \sum_{h=0}^{m(k)-j} \Bigg\{ \binom{m(k)}{j} \frac{m(k)-j}{(m(k)-j-h)!h!} \\ &\times (-1)^h e^{-j\lambda_s T - h\lambda_s T} \\ &\times \gamma\left(m(k)-j, \lambda_s\left(\frac{\Delta}{cw} - jT - hT\right)\right)\Bigg\}, \end{aligned} \tag{28}$$

where $\gamma(b,x) = \int_0^x u^{b-1}e^{-u}du$ is the incomplete Gamma function defined in [33].

Proof: From (27), we have

$$\mathrm{P}(\mathcal{F}) = \mathrm{P}\left(\sum_{j=1}^{m(k)} \min\{\tau_j, T\} < \frac{\Delta}{cw}\right). \tag{29}$$

If $\frac{\Delta}{cw} > m(k)T$, $\mathrm{P}(\mathcal{F}) = 1$. As a result, we will focus on the case where $\frac{\Delta}{cw} \le m(k)T$. Since τ_j's are i.i.d. exponential random variables, $\min\{\tau_j, T\}$ are i.i.d. with distribution

$$\mathrm{P}(\min\{\tau_j, T\} < t) = \begin{cases} 1 - e^{-\lambda_s t} & t < T \\ 1 & t \ge T. \end{cases} \tag{30}$$

$\mathrm{P}(\mathcal{F})$ can be expressed based on (29) as

$$\begin{aligned}\mathrm{P}(\mathcal{F}) = &\sum_{j=0}^{\lfloor \frac{\Delta}{cw}/T \rfloor} \Bigg\{ \binom{m(k)}{j} e^{-j\lambda_s T}\left(1 - e^{-\lambda_s T}\right)^{m(k)-j} \\ &\times \underbrace{\mathrm{P}\left(\widehat{\tau}(m(k)-j) < \frac{\Delta}{cw} - jT\right)}_{\triangleq \psi}\Bigg\}, \end{aligned} \tag{31}$$

where $\lfloor . \rfloor$ is the floor function, $\widehat{\tau}(m(k)-j)$ is the sum of $m(k)-j$ i.i.d random variables with distribution shown in (30) given each of these variables is less than T. Since $\frac{\Delta}{cw} \le m(k)T$, $\left(\frac{\Delta}{cw} - jT\right)/T \le m(k) - j$. From [35], ψ can be derived as

$$\begin{aligned}\psi = &\left(\frac{1}{1 - e^{-\lambda_s T}}\right)^{m(k)-j} \sum_{h=0}^{\lfloor (\frac{\Delta}{cw} - jT)/T \rfloor} \Bigg\{ \frac{m(k)-j}{(m(k)-j-h)!h!} \\ &\times (-1)^h e^{-h\lambda_s T} \gamma\left(m(k)-j, \lambda_s\left(\frac{\Delta}{cw} - jT - hT\right)\right)\Bigg\}. \end{aligned} \tag{32}$$

Plugging (32) into (31), $\mathrm{P}(\mathcal{F})$ can be obtained as shown in Theorem 2. ∎

By definition, $\mathrm{P}(\mathcal{T}_{\vartheta'})$ is the same as in (24). Plugging (24) and (28) into (26), φ can finally be derived.

3) $a = \delta$, $\vartheta' = \delta$: There are two possible ways for the data block to be transferred to the intended direction δ. The first one is that the considered data block is successfully delivered to a potential relay traveling along δ. The other one is that the data transmission fails, but the data-carrying CRV turns to the right direction. The two events corresponding to these two possibilities are denoted as $\overline{\mathcal{F}}$ and $\mathcal{T}_{\mathcal{F},\delta}$, respectively. By definition, $\overline{\mathcal{F}}$ and $\mathcal{T}_{\mathcal{F},\delta}$ will not simultaneously happen. Then, it follows

$$\varphi = \mathrm{P}\left(\overline{\mathcal{F}} \cup \mathcal{T}_{\mathcal{F},\delta}\right) = \mathrm{P}\left(\overline{\mathcal{F}}\right) + \mathrm{P}(\mathcal{T}_{\mathcal{F},\delta}). \tag{33}$$

Since $a(k) = \delta$, we only need to focus on the case where $\eta_\delta(k) = 1$ and $m(k) > 0$. According to the definition of $\overline{\mathcal{F}}$, we have

$$\mathrm{P}\left(\overline{\mathcal{F}}\right) = \mathrm{P}\left(\sum_{j=1}^{m(k)} cw \times \min\{\tau_j, T\} \ge \Delta\right). \tag{34}$$

Clearly from (34), $\mathrm{P}\left(\overline{\mathcal{F}}\right) = 1 - \mathrm{P}(\mathcal{F})$ and thus can be obtained from Theorem 2.

By definition, $\mathcal{T}_{\mathcal{F},\delta} = \mathcal{F} \cap \mathcal{T}_\delta$, where $\mathcal{F}$ and $\mathcal{T}_\delta$ share the same definition as in (26). With the independence between $\mathcal{F}$ and $\mathcal{T}_\delta$, $\mathrm{P}(\mathcal{T}_{\mathcal{F},\delta})$ can be expressed as

$$\mathrm{P}(\mathcal{T}_{\mathcal{F},\delta}) = \mathrm{P}(\mathcal{F})\,\mathrm{P}(\mathcal{T}_\delta). \tag{35}$$

The expression for $\mathrm{P}(\mathcal{F})$ is the same as that derived in the case with $a(k) = \delta$ and $\vartheta' \ne \delta$. $\mathrm{P}(\mathcal{T}_\delta)$ is the same as φ shown in (24) with $\vartheta' = \delta$. Together with (33), (34), and (35), φ can be obtained accordingly.

With the results in IV.C and IV.D, the state transition probabilities can finally be obtained by plugging the expressions of ϕ and φ into (3).

E. Objective and Reward

As aforementioned, the SSP tries its best to deliver the considered data block to the intended destination in K decision epochs. To put in a mathematical way, the SSP should find a policy $g*$ which maximizes $\mathcal{R}(g) = \mathrm{E}\left[\sum_{k=1}^{K} R(s(k), s(k+1), a(k))\right]$, where $a(k)$'s are actions selected according to the policy g, $R(s(k), s(k+1), a(k))$ is the reward received from the kth decision epoch with

$$R(s(k), s(k+1), a(k)) = \begin{cases} R & \mathcal{I}(k) \neq n^d, \mathcal{I}(k+1) = n^d, 1 \leq k \leq K, \\ 0 & Otherwise \end{cases} \tag{36}$$

where R is the reward obtained by successfully delivering the considered data block to the destination. From (36), obtaining $g* = \arg\max_g \{\mathcal{R}(g)\}$ is equivalent to finding a policy such that the probability that the considered data block is successfully delivered to the destination is maximized.

With the state transition probabilities and the reward function presented above, the SSP can obtain the data routing decisions to help CRVs select data forwarding actions by solving the formulated MDP with dynamic programming [36]. ***Remark***: From [36], dynamic programming solves the formulated MDP through backward induction. Namely, it starts from the last decision epoch and continues backwards to obtain the optimal action at each state during each decision epoch and the corresponding optimal total expected reward, which finally gives us the optimal policy. Noticing that the formulated MDP is a finite horizon problem with K decision epochs, dynamic programming solves the formulated MDP in K steps. In the $(K-k+1)$th step, the optimal action at each state during the kth decision epoch and the corresponding expected optimal total reward are derived based on the optimal total expected reward of the $(k+1)$th decision epoch obtained during the $(K-k)$th step. Notice that the formulated MDP has $\Theta(n_{\mathcal{I}} n_m)$ states, where $n_{\mathcal{I}}$ is the number of intersections considered in the formulation, and n_m is the maximum number of frequency bands in the considered area, i.e., $n_m = \max_n M_n$, M_n is the number of frequency bands at the nth intersection considered in the formulation. Then, to solve the formulated MDP, each step of dynamic programming will cost $\Theta\left(n_a n_{\mathcal{I}}^2\ n_m^2\right)$ time, where n_a is the number of available actions. Given K is the total number of decision epochs, we have $n_{\mathcal{I}} = \mathrm{O}\left(K^2\right)$. From Section IV.A, $n_a = 5$. Thus, dynamic programming will solve the formulated MDP in $\mathrm{O}\left(K^5\ n_m^2\right)$ time, which implies that the time consumed by dynamic programming to solve the formulated MDP is affected by the total number of decision epochs and the total number of frequency bands in the considered system.

V. Performance Evaluation

In this section, extensive simulation results are presented to evaluate the performance of the obtained data routing decisions and the effectiveness of the V-CCHN in data transportation.

A. Considered Scenarios & Parameter Settings

In this section, we consider a road network similar to that in Fig. 2, but, with 21 rows, the east-westward roads, and 21 columns, the north-southward roads. Intersection D locates at the same column as intersection $\mathcal{S}$ and is κ rows above intersection $\mathcal{S}$. CRVs traverse each road segment in a time unit and the considered data block needs to be delivered in K times units. Since the first data routing decision will be made at intersection $\mathcal{S}$, $\kappa \leq K$. Otherwise, it is impossible for the SSP to deliver the considered data block to intersection D after K decision epochs. If the considered data block is delivered to intersection D through K decision epochs or less, a reward of $R = 500$ will be received. The set $\overline{\Theta}$ is identified by the following procedure

1) Construct a graph with intersections as vertices and there is an edge between two vertices if the corresponding intersections are adjacent.
2) Calculate the length of the shortest paths from the vertex associated with intersection $\mathcal{S}$ to other vertices. $\mathcal{L}_v^{\mathcal{S}}$ is the length of the shortest path to vertex v.
3) Calculate the length of the shortest paths from the vertex associated with intersection D to other vertices. $\mathcal{L}_v^D$ is the length of the shortest path to vertex v.
4) Include the intersection corresponding to vertex v to $\overline{\Theta}$ if $\mathcal{L}_v^{\mathcal{S}} + \mathcal{L}_v^D > K$.

The set Θ consists of intersections other than those in $\overline{\Theta}$. Unless explicitly stated, we assume that no obstacle for data transmissions exists at intersections to highlight the effects of spectrum availability on data routing decisions. The average numbers of potential relays at different intersections along different directions are λ_V, i.e., $\lambda_n^N = \lambda_n^S = \lambda_n^W = \lambda_n^E = \lambda_V$, $\forall n \in \Theta_{\mathcal{I}}$. Potential relays traveling along the different directions are occupied with the same probability p, i.e., $p_n^N = p_n^S = p_n^W = p_n^E = p$. At intersections, the data-carrying CRV will keep current direction, turn left, turn right, and turn around with probability, $p_1 = 0.36$, $p_2 = 0.3$, $p_3 = 0.3$, and $p_4 = 0.04$, respectively. λ in (8) is set to 0.5. At each decision epoch, each band is found to be available with probability ρ, i.e., $\rho_n = \rho$, $\forall n \in \Theta_{\mathcal{I}}$. If found available, a band will remain available for 7 s on average, i.e., $\lambda_s = 7$. Each band has a bandwidth of 10 MHz. The links between the data-carrying CRVs and the relaying CRVs can achieve a spectral efficiency of 2 $bit/s/Hz$ and can last for $T = 10$ s at different intersections. When the considered data block arrives at intersection $\mathcal{S}$, 3 harvested bands are available, and potential relays traveling along south, west, and east are available to receive this considered data block. The data-carrying CRV approaches intersection $\mathcal{S}$ from south. Similar to [27], the initial distribution of vehicles on the considered road network is generated according to the steady state distribution and thus set-up time is not used. The following results are obtained from 10000 rounds of simulations.

B. Simulation Results and Discussions

In Fig. 3, we examine the impacts of Δ, the size of the considered data block, and the level of contention at intersections on the probability of successful delivery, i.e., the probability that the considered data block is delivered to the destination after $K = 9$ decision epochs. Intersection D is $\kappa = 5$ rows above intersection $\mathcal{S}$. The level of contention at the intersections is represented by p, the probability that a potential relay is involved with the transmissions of previously arrived data blocks. The total number of bands is $M_n = 10$, $\forall n \in \Theta_{\mathcal{I}}$ and each band is available with probability $\rho = 0.5$.

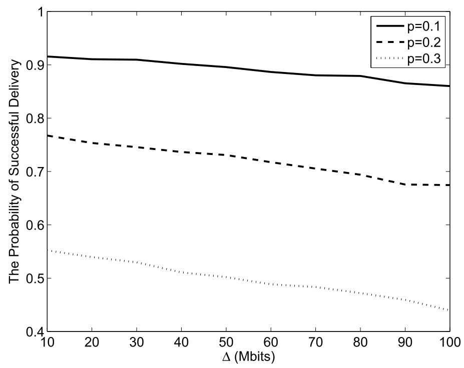

Fig. 3. The probability of successful delivery v.s. the size of the considered data block and the level of contention at each intersection.

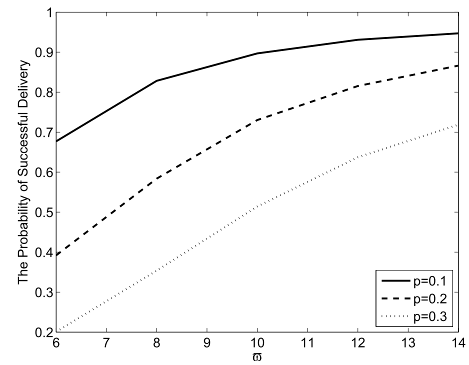

Fig. 4. The probability of successful delivery v.s. the number of frequency bands, ϖ.

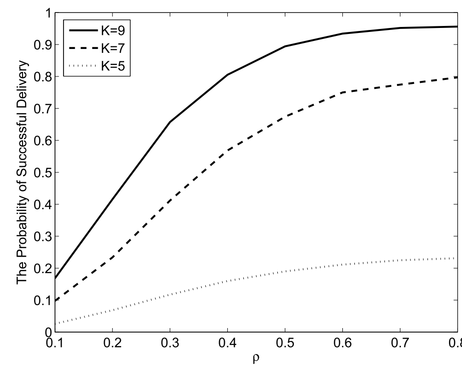

Fig. 5. The probability of successful delivery v.s. the activities of licensed/unlicensed users and the delay tolerance of the considered data block.

λ_V is set to 2. As shown in Fig. 3, the probability of successful delivery decreases with Δ. Under the same level of resource provisioning, it becomes more difficult for the data-carrying CRV to transfer the considered data block to a relaying CRV when the size of this data block increases. Due to the increase in Δ, the considered data block often cannot be routed to the designated directions, which leads to a decrease in the probability of successful delivery. Similarly, the probability of successful delivery decreases with p increasing, as shown in Fig. 3, since transferring the considered data block from the data-carrying CRV to the relaying CRV will get harder due to the higher level of contention at each intersection. Besides, it can be observed from Fig. 3 that considerable probability of successful delivery can be achieved if the SSP can provide enough resources to facilitate the transportation of the considered data block (the case when $p = 0.1$). This is just the reason why we propose to employ CR technologies in our V-CCHN to harvest the wide range of under-utilized spectrum resources. With enough resources acquired, our V-CCHN is certainly capable of providing the envisioned data transportation services. This observation can be further strengthened by the results as shown in Fig. 4.

In Fig. 4, we investigate how the probability of successful delivery varies with the number of frequency bands in the system. For illustration purpose, we set $M_n = \varpi$, $\forall n \in \Theta_{\mathcal{I}}$. The parameter settings are the same as those in Fig. 3, other than $\Delta = 50\ Mbits$. The results show that the SSP can achieve a higher probability of successful delivery if there are more frequency bands for it to exploit. By comparing the results with different p, i.e., different levels of contention, it is clear that the SSP can effectively overcome the contention at intersections by utilizing under-utilized licensed/unlicensed bands via CR technologies. With enough spectrum resources, the considered data block is more likely to be transferred to desired directions. Thus, the SSP can more efficiently address the uncertainty in the mobility of CRVs and facilitate the envisioned data transportation services, which further validates our V-CCHN.

Fig. 5 presents how the activities of licensed/unlicensed users, signified by ρ, and the delay tolerance of the considered data block, represented by K, affect the probability of successful delivery. The parameter settings are the same as those in Fig. 4 and the only differences are $p = 0.1$ and $M_n = 10, \forall n \in \Theta_{\mathcal{I}}$. From Fig. 5, a higher value of ρ improves the probability of successful delivery. Since the bands will be found available with higher probability under a higher value of ρ, the SSP could acquire more available bands to enable the transferring of the considered data block to relaying CRVs. With more spectrum resources, the considered data block could be successfully delivered to the destination with a higher probability, as demonstrated in Fig. 4. Fig. 5 also indicates that the delay tolerance of the considered data block, K, has significant impacts on the probability of successful delivery. As shown in Fig. 3 and Fig. 4, the data delivery in the V-CCHN is subject to the impacts of the activities of licensed/unlicensed users and the mobility of CRVs, which introduces uncertainty to data routing at intersections. Because of these dynamics in the V-CCHN, the considered data block might not be delivered to the desired directions at some intersections. If the data block can tolerate longer delay, the SSP will have more opportunities to make corrections on the routing of this data block and be more capable of handling the dynamics inherent in the V-CCHN, which improves the probability of successful delivery. This is also the reason why the V-CCHN is targeted to handle delay-tolerant data traffic.

A similar conclusion can be drawn from Fig. 6 where the relationship between the probability of successful delivery and the distance between the source, intersection $\mathcal{S}$, and the destination, intersection D, is presented. For ease of illustration, intersection D stays in the same column as intersection $\mathcal{S}$, but, locates κ rows above intersection $\mathcal{S}$. The parameter

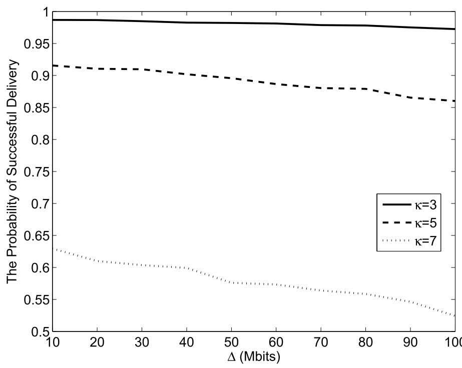

Fig. 6. The probability of successful delivery v.s. the distance between the source and the destination.

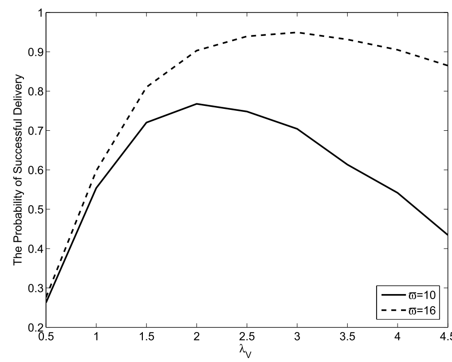

Fig. 7. The impacts of the density of vehicles on the probability of successful delivery.

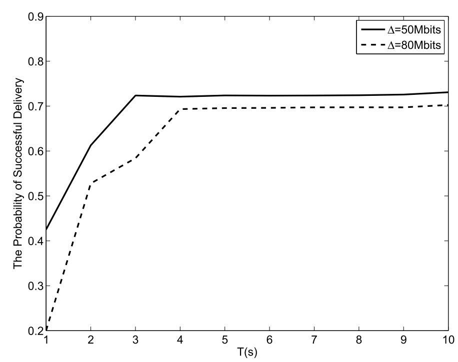

Fig. 8. The impacts of vehicular mobility on the probability of successful delivery.

settings are the same as Fig. 3 other than $p = 0.1$. The results show that the probability of successful delivery decreases with κ increasing. Intuitively, the considered data block will travel a longer distance when κ increases. Once such distance gets longer, as aforementioned, more randomness will be introduced to the data delivery process due to the dynamics inherent in the V-CCHN. In this case, the SSP needs more rooms to deal with these dynamics/randomness and ensure successful delivery, as shown in Fig. 5. This explains why, in Fig. 6, the probability of successful delivery decreases as κ increases since the curves are obtained under the same level of delay tolerance. Combining the results in Fig. 5 and Fig. 6, we can conclude that the V-CCHN can effectively handle the transportation of delay-tolerant data traffic once the delivered data can tolerate long enough delays.

Given that the density of vehicles is reflected through the average number of potential relays λ_V, we investigate its impacts on the data delivery in the V-CCHN by studying how the probability of successful delivery varies with λ_V. The results are shown in Fig. 7. The parameter settings are the same as those in Fig. 4 except $\Delta = 10\ Mbits$. ϖ is the total number of frequency bands at each intersection. Clearly from Fig. 7, the probability of successful delivery first increases with λ_V, which matches our intuition since the probability of successful delivery is initially limited by the availability of potential relays. Additionally, it can be observed from Fig. 7 that, after λ_V exceeds a certain value, further increase in λ_V will result in a decrease in the probability of successful delivery. Notice that increasing λ_V will not only increase the availability of potential relays but also increase the contention at intersections. Because of this increased contention, the probability of successful delivery decreases after λ_V exceeds a certain value.

In Fig. 8, we employ the contact duration T between vehicles as an example to study how vehicular mobility affects the performance of data delivery in the V-CCHN. The results are shown in Fig. 8. The parameter settings are the same as those in Fig. 3. From Fig. 8, when T is small, i.e., in the highly dynamic scenario, the probability of successful delivery is limited by the short contact duration between vehicles. It can be observed from Fig. 8 that the probability of successful delivery reaches a plateau after T reaches certain value. In this case, we should look for other ways, such as trying to harvest more spectrum resources, to improve the probability of successful delivery, since it is no longer limited by the contact duration between vehicles.

In Fig. 9, we study how the presence of obstacles at intersections affects the probability of successful delivery. The parameter settings are the same as those in Fig. 3. The only difference is that, considering the shadowing effects due to obstacles, the links between two vehicles moving along perpendicular directions have a spectral efficiency of $0.5\ bit/s/Hz$. From Fig. 9, the shadowing effects resulted from obstacles have significant impacts on data delivery, particularly when the size of the carried data block is large. To address this issue, we could deploy wireless routers/relays at pivotal intersections to help with data transmissions, and the deployment of these routers/relays will be the focus of one of our future works.

To evaluate the effectiveness of our spectrum-aware data transportation scheme, we compare its data delivery performance with the performance of the greedy perimeter stateless routing (GPSR), a widely adopted benchmark scheme, and the performance of the Minimum Delay Routing Algorithm (MDRA) introduced in [27]. GPSR uses greedy forwarding to forward data to the vehicle which leads to the maximum progress towards the destination [25]. In MDRA, the considered block is always transferred towards the direction which leads to the most significant improvement in the delivery delay when wireless communications are assumed to be perfect. When compared with GPSR, MDRA explicitly considers the

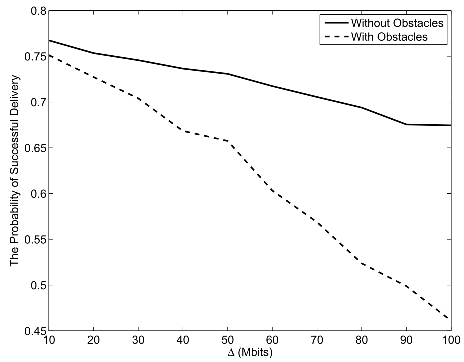

Fig. 9. The impacts of the presence of obstacles on the probability of successful delivery.

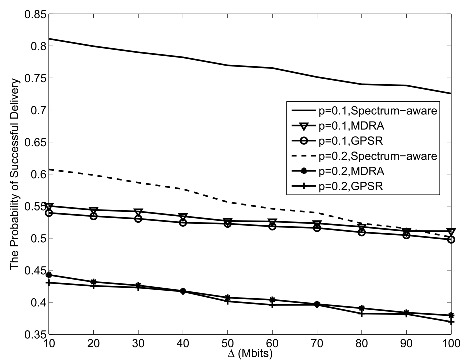

Fig. 10. The probability of successful delivery under different schemes. The proposed scheme is labeled as "Spectrum-aware".

Fig. 11. The considered area in Shanghai (latitude: from 31.2133 to 31.2217, longitude: from 121.4684 to 121.4768).

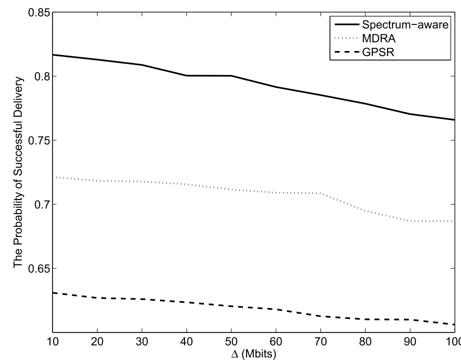

Fig. 12. The probability of successful delivery based on real trace driven simulations. The proposed scheme is labeled as "Spectrum-aware".

availability of vehicles at subsequent intersections. To adapt GPSR and MDRA to our scenario, we adopt them for decision making at intersections and add buffers to GPSR as illustrated in [14] and [25]. The results are shown in Fig. 10. The parameter settings are the same as in Fig. 3 except that no harvested band is available at intersections (x, y) $(x \in \{-1, -2\}$ and $y \in \{0, 1, 2\})$, where intersection $(-1, 1)$ is one row above the 0th row and one column to the right of the 0th column. It can be observed from Fig. 10 that our spectrum-aware scheme can significantly improve the probability of successful delivery when compared with both GPSR and MDRA. The superiority of our scheme is inherent in its spectrum-aware design which judiciously exploits information on contentions and the activities of licensed/unlicensed users to route data blocks in order to circumvent intersections lacking spectrum resources and facilitate efficiently data delivery. Thus, when compared with existing schemes, the proposed spectrum-aware scheme can more efficiently support the data transportation in the V-CCHN.

Finally, we further evaluate the effectiveness of our scheme based on real trace driven simulations where the turning probabilities of CRVs at intersections are obtained through real traces. Specifically, we consider an area in Shanghai as shown in Fig. 11 (latitude: from 31.2133 to 31.2217, longitude: from 121.4684 to 121.4768) and utilize the real traces collected from 4316 taxies in Shanghai between $10:00$am and $1:00$pm on Feb 20, 2007 (available at http://www.cse.ust.hk/scrg) to obtain the turning probabilities [37]. In each round of simulation, a data block, to be delivered to the intersection with latitude 31.2165 and longitude 121.4708, is picked up at the intersection with latitude 31.2176 and longitude 121.4742. The other settings, such as the availability of harvested bands and that of potential relays, are the same as those in Fig. 3, except that $K = 6$, $p = 0.1$, and no harvested band is available at the intersection with latitude 31.2171 and longitude 121.4724. The results are shown in Fig. 12, where the performance of our scheme is compared with the performance of GPSR and the performance of MDRA. From Fig. 12, our spectrum-aware data transportation scheme outperforms both GPSR and MDRA in the real trace driven simulation. Moreover, as shown in Fig. 12, based on our proposed scheme, considerable probability of successful delivery can be achieved in the considered real trace driven simulation, which further demonstrates the effectiveness of our proposed approach.

VI. Conclusion

In this paper, we design a spectrum-aware data transportation scheme for our recently proposed V-CCHN architecture by formulating the data delivery process as a Markov decision process. Through extensive simulations, we demonstrate that the obtained data transportation scheme can effectively utilize the spectrum opportunity and mobility opportunity in the V-CCHN for data transportation. This implies that, with properly designed data transportation schemes, our V-CCHN offers us a very promising alternative to handling the soaring wireless data traffic in the incoming era of smart cities. Thus, we hope this work will trigger more research activities and efforts to explore and further develop such an intelligent data transportation network.

References

[1] T. Han *et al.*, "5G converged cell-less communications in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 44–50, Mar. 2017.

[2] 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.

[3] S. Djahel, R. Doolan, G.-M. Muntean, and J. Murphy, "A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 125–151, 1st Quart., 2015.

[4] N. Cheng *et al.*,, "Vehicle-assisted device-to-device data delivery for smart grid," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2325–2340, Apr. 2016.

[5] J. M. Batalla *et al.*, "Efficient media streaming with collaborative terminals for the smart city environment," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 98–104, Jan. 2017.

[6] C.-X. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[7] I. Yaqoob *et al.*, "Enabling communication technologies for smart cities," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 112–120, Jan. 2017.

[8] D. Mazza, D. Tarchi, and G. E. Corazza, "A unified urban mobile cloud computing offloading mechanism for smart cities," *IEEE Wireless Commun.*, vol. 55, no. 3, pp. 30–37, Mar. 2017.

[9] 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.

[10] F. Malandrino, C. Casetti, C.-F. Chiasserini, and M. Fiore, "Optimal content downloading in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1377–1391, Jul. 2013.

[11] H. Yao *et al.*, "Opportunistic offloading of deadline-constrained bulk cellular traffic in vehicular DTNs," *IEEE Trans. Comput.*, vol. 64, no. 12, pp. 3515–3527, Dec. 2015.

[12] F. Xu *et al.*, "Utilizing shared vehicle trajectories for data forwarding in vehicular networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 441–445.

[13] K. Abboud, H. A. Omar, and W. Zhuang, "Interworking of DSRC and cellular network technologies for V2X communications: A survey," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9457–9470, Dec. 2016.

[14] S. M. Tornell, C. T. Calafate, J.-C. Cano, and P. Manzoni, "DTN protocols for vehicular networks: An application oriented overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 868–887, 2nd Quart., 2015.

[15] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Efficient routing in intermittently connected mobile networks: The multiple-copy case," *IEEE/ACM Trans. Netw.*, vol. 16, no. 1, pp. 77–90, Feb. 2008.

[16] H. Zhu, S. Chang, M. Li, K. Naik, and S. Shen, "Exploiting temporal dependency for opportunistic forwarding in urban vehicular networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2192–2200.

[17] H. Zhu *et al.*, "ZOOM: Scaling the mobility for fast opportunistic forwarding in vehicular networks," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 2832–2840.

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

[19] Y. Zhu, Y. Wu, and B. Li, "Trajectory improves data delivery in urban vehicular networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 4, pp. 1089–1100, Apr. 2014.

[20] Y. Wu, Y. Zhu, H. Zhu, and B. Li, "CCR: Capacity-constrained replication for data delivery in vehicular networks," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 2580–2588.

[21] T. Darwish and K. A. Bakar, "Lightweight intersection-based traffic aware routing in Urban vehicular networks," *Comput. Commun.*, vol. 87, pp. 60–75, Aug. 2016.

[22] X. Zhang, X. Cao, L. Yan, and D. K. Sung, "A street-centric opportunistic routing protocol based on link correlation for urban vanets," *IEEE Trans. Mobile Comput.*, vol. 15, no. 7, pp. 1586–1599, Jul. 2016.

[23] G. Li, L. Boukhatem, and J. Wu, "Adaptive quality-of-service-based routing for vehicular ad hoc networks with ant colony optimization," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3249–3264, Apr. 2017.

[24] M. A. Togou, A. Hafid, and L. Khoukhi, "SCRP: Stable CDS-based routing protocol for urban vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 5, pp. 1298–1307, May 2016.

[25] 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.

[26] T. S. J. Darwish, K. A. Bakar, and K. Haseeb, "Reliable intersection-based traffic aware routing protocol for urban areas vehicular ad hoc networks," *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 1, pp. 60–73, Jan. 2018.

[27] 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.

[28] H. Ding *et al.*, "Cognitive capacity harvesting networks: Architectural evolution toward future cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1902–1923, 3rd Quart., 2017.

[29] H. Ding *et al.*, "Session-based cooperation in cognitive radio networks: A network-level approach," *IEEE/ACM Trans. Netw.*, vol. 26, no. 2, pp. 685–698, Apr. 2018.

[30] P. Si, Y. He, H. Yao, R. Yang, and Y. Zhang, "DaVe: Offloading delay-tolerant data traffic to connected vehicle networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 3941–3953, Jun. 2016.

[31] Y. Long, H. Li, H. Yue, M. Pan, and Y. Fang, "SUM: Spectrum utilization maximization in energy-constrained cooperative cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 11, pp. 2105–2116, Nov. 2014.

[32] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1975.

[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 7th ed. Amsterdam, The Netherlands: Elsevier, 2007.

[34] M. Pan, C. Zhang, P. Li, and Y. Fang, "Spectrum harvesting and sharing in multi-hop CRNs under uncertain spectrum supply," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 369–378, Feb. 2012.

[35] H. Ding, C. Zhang, B. Lorenzo, and Y. Fang, "Access point recruitment in a vehicular cognitive capability harvesting network: How much data can be uploaded?" *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6438–6445, Jul. 2018.

[36] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, 3rd ed. Belmont, MA, USA: Athena Scientific, 2005.

[37] J. He, L. Cai, P. Cheng, and J. Pan, "Delay minimization for data dissemination in large-scale VANETs with buses and taxis," *IEEE Trans. Mobile Comput.*, vol. 15, no. 8, pp. 1939–1950, Aug. 2016.

Haichuan Ding received the B.E. and M.S. degrees in electrical engineering from the Beijing Institute of Technology, Beijing, China. He is currently pursuing the Ph.D. degree at the University of Florida, Gainesville, FL, USA. His research interests include cognitive radio networks, vehicular networks, and security and privacy in distributed systems.

Xuanheng Li is currently a Lecturer with the School of Information and Communication Engineering, Dalian University of Technology. His research interests include cognitive radio networks and Internet of Things.

Ying Cai is currently a Full Professor with Beijing Information Science & Technology University, Beijing. Her current research interests include cyber security and wireless networks.

Beatriz Lorenzo is currently a Senior Researcher with the Atlantic Research Center for Information and Communication Technologies, University of Vigo, Vigo, Spain. Her research interests include design and optimization of wireless networks and complex networks.

Yuguang Fang is currently a Professor with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. His research interests include wireless networks, privacy and security.