IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 25, 2018, accepted October 5, 2018, date of publication October 11, 2018, date of current version November 14, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2875566

Geographic Segmented Opportunistic Routing in Cognitive Radio Ad Hoc Networks Using Network Coding

XING TANG[1], (Member, IEEE), JUNWEI ZHOU[1], (Member, IEEE), SHENGWU XIONG[1], JING WANG[2], AND KUNXIAO ZHOU[3]

[1]School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China
[2]Computer School of Wuhan University, Wuhan 430072, China
[3]School of Computer Science and Network Security, Dongguan University Technology, Dongguan 523808, China

Corresponding author: Junwei Zhou (junweizhou@msn.com).

This work was supported in part by the National Natural Science Foundation of China under Grant 61601337 and Grant 61771354, in part by the Hubei Provincial Natural Science Foundation of China under Grant 2017CFB593, and in part by the Fundamental Research Funds for the Central Universities under Grant WUT:2017II14XZ.

ABSTRACT In cognitive radio ad hoc networks (CRAHNs), the secondary users' links are interrupted by the arrival of primary users, which leads to a significant increase of the number of transmissions per packet. Since network coding and opportunistic routing can reduce the number of transmissions over unreliable wireless links, they are more suitable for CRAHNs. Network coding-based opportunistic routing has been well studied in previous works for traditional ad hoc networks. However, existing approaches have limitations to handle CRAHNs. First, the routing decisions of these methods aim to find one optimal forwarding set for the whole path from the source to the destination. If the channel conditions change dynamically, the pre-selected forwarding set may become unavailable or non-optimal. Second, their coding schemes are not efficient enough to cope with the dynamic spectrum variation in CRAHNs. In this paper, we attempt to overcome these limitations and propose a network-coding-based geographic segmented opportunistic routing scheme for CRAHNs, by fully embracing its characteristics. In our scheme, the whole path from the source to the destination is cut down into several smaller opportunistic route segments, where the packets are transmitted through multiple segments based on a step-by-step forwarding procedure until all the packets are delivered to the destination. Since our scheme utilizes only local spectrum opportunities, topology information, and geometric conditions to compute the forwarding set for each short-term opportunistic route segment, it can better adapt to dynamic spectrum environments and changing network topologies in CRAHNs. Furthermore, we construct a coding graph to show that our coding problem is a reduction from the maximum clique problem, and we propose an efficient network coding strategy to solve it. Simulation results show that our scheme can achieve a considerable performance, compared with the latest opportunistic routing protocols designed for CRAHNs.

INDEX TERMS Cognitive radio, ad hoc networks, routing protocols, geographic opportunistic routing.

I. INTRODUCTION

COGNITIVE Radio (CR) networks have emerged as a promising technology to improve the spectral efficiency of the existing radio spectrum [1]. In CR networks, secondary users (SUs) can sense and access the current wireless spectrum opportunistically. Since the nodes have different spectrum opportunities, the available channels of SUs are different and changing dynamically. In CR ad hoc networks (CRAHNs), the lack of infrastructure support and the instability make routing in such networks a challenging issue [2].

Various routing protocols have been proposed for CRAHNs [3]. These efforts focus on designing new routing metrics by considering the dynamic spectrum characteristics of CRAHNs, and they attempt to compute the best end-to-end path for data transmission. Nevertheless, in practice, establishing such a route for a whole flow is usually not

possible due to the instability and unpredictability of channel conditions in CRAHNs. Moreover, most existing schemes rely on a computed routing path to forward the packets hop by hop along the path. Since the link conditions in CRAHNs are usually poor (e.g., low delivery ratio or poor reliability), the data transmission along a pre-fixed path would incur a lot of retransmissions.

Opportunistic routing [4] does not fix the route before the packet transmission. It can reduce the number of transmissions over unreliable and unpredictable wireless links by taking advantage of the broadcast nature of the wireless medium. Besides, geographic routing [5] does not require the establishment and maintenance of the whole path from the source to the destination, and thus it is an attractive option in the case of dynamic links. Recently, some initial work for opportunistic routing in CR environments can be found in [6]–[8]. The majority of these studies show that geographic routing can work together with opportunistic routing (geographic opportunistic routing) to improve the data delivery performance. Most of these solutions aim to find one or more paths for each long-term connection between one source and one destination. However, if the channel conditions change dynamically, the pre-selected candidates and channels may become an unavailable and non-optimal solution, which brings a significant challenge.

In this paper, we mainly focus on geographic opportunistic routing in CRAHNs, and our objective is to minimize the average number of transmissions required to deliver a packet from the source to the destination. Different from the existing geographic opportunistic routing solutions that calculate the optimal forwarding set for the whole path from the source to the destination, our geographic segmented opportunistic routing (GSOR) scheme divides the end-to-end path into several smaller opportunistic route segments and transmits the packets based on a step-by-step forwarding procedure. Since GSOR utilizes only local spectrum opportunities, topology information and geometric conditions to compute the forwarding set for each short-term opportunistic route segment, it can better adapt to dynamic spectrum environments and changing network topologies.

Network coding [9] is another emerging technology that can further reduce the number of transmissions by exploiting the broadcast nature of wireless medium. In this paper, we propose to integrate network coding with GSOR in CRAHNs, which is motivated by the following considerations. First, since network coding can reduce the number of transmissions, combining network coding with GSOR can be beneficial to the objective of this work. Second, as considered in [10], opportunistic routing can improve the performance of network coding in lossy networks. Since opportunistic routing scatters the packets of one unicast flow over multiple paths, it brings more coding opportunities without forcing flows to intersect at some rendezvous nodes. Finally, it is beneficial to apply network coding in CRAHNs. The link conditions in CRAHNs are unsteady, and the communication interruptions may cause a large number of retransmissions. Many studies [11], [12] confirm that network coding can decrease the number of retransmissions over unreliable wireless links, thus it can improve the routing performance in CRAHNs.

Network coding based opportunistic routing has been extensively studied in traditional ad hoc networks. Nevertheless, these techniques have some limitations in CRAHNs. As mentioned before, the routing solutions for traditional networks are not suitable for CR environments. Moreover, their coding solutions are not efficient enough to cope with the dynamic spectrum variation in practical CR contexts. Recently, several strategies have been proposed to incorporate network coding into routing protocols for CR networks [13], which can increase the channel availability [14], maximize the throughput [15], [16], reduce the end-to-end delay violation probability [17] and minimize the total transmission cost [18]. However, to the best of our knowledge, no prior studies have been reported on network coding based geographic opportunistic routing with the objective of minimizing the average number of transmissions per packet.

In summary, the major contributions of this paper are listed as follow.

- We propose a GSOR mechanism in which the whole path from the source to the destination is cut down into several opportunistic route segments, and the packets are transmitted through these segments based on a step-by-step forward procedure. Specifically, the intermediate destination node and relay nodes of each segment are selected independently and based on only local channel and topology information, which can quickly adapt to the dynamic changing CR environments.
- We design a local metric to characterize the packet delivery ratio for each opportunistic route segment, and theoretically analyze the properties of the metric. With the help of the analysis results, we propose an optimal relay selection algorithm to achieve the upper bound of the metric. To reduce the computation complexity, we also propose a sub-optimal solution. Moreover, we develop a greedy temporary destination selection algorithm to balance the trade-off between the performance and the distance advancement.
- We construct a coding graph to show that our coding problem is a reduction from the maximum clique problem and propose a heuristic greedy approach to solve it efficiently. Moreover, we discuss the implementation details, such as how to obtain reception information and how to control the data receiving/sending processes.

Note that this paper is an extension of its conference version in [19]. The new materials in the current work are the following: i) In subsection IV-B, we design a new metric to measure the packet delivery ratio of one opportunistic route segment, and investigate the corresponding properties and the lower bound of the metric through theoretical analyses. ii) In subsection IV-C, we prove that the upper bound can be achieved by a greedy relay selection algorithm

we developed. We also propose a sub-optimal solution to balance the trade-off between the performance and the complexity. iii) In subsection IV-E, we modify the coding method by applying a new metric ''effective degree''. Besides, we describe the implementation details of our scheme. iv) In section V, more results are reported to show the end-to-end delay and the energy consumption of our scheme, compared with more recently developed methods.

The rest of the paper is organized as follows. Related work is discussed in section II. In section III, we describe the system model and problem statement. We propose the network coding based GSOR scheme in section IV. Performance evaluation is provided in section V. Section VI concludes this paper and points out future research directions.

II. RELATED WORK

Various routing protocols have been proposed for CR networks. We can categorize them into two main classes depending on the awareness of spectrum information. The first approaches [20]–[22] are based on full spectrum knowledge with the assumption that the spectrum information is available to all the nodes in the network. The primary task of these approaches is to make routing decision without spectrum assessment. The second category includes the routing schemes based on local spectrum knowledge. Such methods usually apply various cross-layer optimizations, joint considering link scheduling, power control and clustering, to minimize the delay [6], [7], maximize the network throughput [23] and the secondary average service rate [24], minimize the energy consumption [25], minimize the packet delivery ratio [26] and enhance the stability of the network [27]. However, most work above set up the whole path for routing without utilizing the broadcast nature of wireless networks.

Extreme Opportunistic Routing (ExOR) was first proposed in [4], which takes advantage of broadcast nature of wireless networks. Some neighboring nodes can overhear the transmission when a node transmits packets. A set of prioritized candidate forwarding nodes is chosen to forward the packets collaboratively. Recently, many researchers are interested in applying opportunistic routing technique to ad hoc networks [28], wireless sensor networks [29], underwater acoustic sensor networks [30], mobile social networks [31], wireless local area networks [32] and vehicular cyber-physical system [33]. Nevertheless, in CR environments, the above articles have suffered from several limitations due to the dynamic nature of the available spectrum.

More recently, some initial work for multi-path opportunistic routing in CR environments can be found in [6]–[8]. In [6], Liu *et al.* propose a novel metric and a heuristic relay selection algorithm to achieve the lowest delay. The authors in [7] provide a complete analysis for the transmission delay over underlay CRAHNs and propose spectrum-map-empowered opportunistic routing algorithms to establish the reliable end-to-end transportation. In [8], the authors design a new cognitive anypath routing metric and propose a polynomial-time routing algorithm to find the best forwarding relays. Note that in our scheme, the path from the source to the destination is cut down into several opportunistic route segments or paths, which is not addressed in the previous articles.

The concept of network coding is first proposed by Ahlswede *et al.* [9] for tackling the multicast issue. Since then Li *et al.* [34] show that linear codes are sufficient to achieve the maximum capacity bounds. Moreover, recent studies (e.g., [18], [35], [36]) show that network coding can also be applied to improve the performance of CR networks [13]. To maximize the multicast rate under multiple constraints, Qu *et al.* in [18] formulate the problem as a chance-constrained program, then apply the Lagrangian relaxation-based optimization to propose an efficient solution for it. In [35], the authors aim to minimize the mean square error at the destinations and design an asymmetric network coding solution for data transmission between the BS and the users. Liang *et al.* in [36] develop an adaptive dynamic network coding scheme for the cooperative communication between PUs (primary users) and SUs by adopting the powerful turbo trellis coded modulation (TTCM). Nonetheless, the idea of combing network coding with opportunistic routing in CR environments was not addressed in the literature.

A lot of work [37]–[39] has been done on the topic of network coding in opportunistic routing for wireless networks. MAC-Independent Opportunistic Routing and Encoding Protocol (MORE) [37] is an extension of the ExOR method using intra-low network coding, where the source mixes the packets by a random linear coding combination. Then, the solutions proposed in [38] and [39] improve the performance of MORE and alleviate MORE's ''stop-and-wait'' nature between consecutive batches. Unfortunately, these earlier studies do not consider the scenario of opportunistic spectrum access. Recently, Zhong *et al.* in [15] start to investigate the coding based opportunistic routing schemes for CR environments. This work focuses on designing a proper metric to select the relay nodes for the whole path from only one source to one destination, combining the existing coding method proposed in [40], without considering geographic location information of the nodes. However, most of these studies consider intra-flow network coding or ignore CR environments. In this work, we attempt to integrate inter-flow network coding with geographic opportunistic routing in CRAHNs. To better utilize the benefit of network coding, in our approach, the whole path is cut down into several independent opportunistic route segments that are associated with multiple intermediate source-destination pairs and forwarding relays, which can bring more coding opportunities and better adapt to the dynamic spectrum environments. Especially for this multi-source multi-destination case, we develop a heuristic coding method and integrate it with the proposed GSOR scheme.

III. SYSTEM MODEL AND PROBLEM DEFINITION

A. SYSTEM MODEL

We consider a CRAHN consisting of infrastructure-supported primary networks and ad hoc secondary networks in the same geographic region. In each cell of the primary network, a single central node (e.g., TV broadcaster or base station) is referred to as a primary transmitter (PT), serving one or multiple primary receivers. The licensed spectrum is divided into a non-empty set of non-overlapping channels, denoted by $\mathcal{K}$. We assume that PTs are geographically distributed following a point Poisson process with different density $\rho_{p,i}$ for each channel $c_j \in \mathcal{K}$. The PTs are assumed to be separated by the interference range, which is about twice the transmission range, to avoid interferences [41]. According to Marten Hardcore Process [42], this distribution can be realized by eliminating overlapping PTs from the original Poisson process model. When a PT is transmitting data on a specific channel, the SUs in the interference range of the PT are not permitted to utilize the same channel simultaneously. In this model, we also assume that PUs who belong to the same PT have the same temporal channel-usage statistics.

In the secondary network, a number of N SUs (nodes) are uniformly distributed with the density of λ_s, and each node n_i $(1 \leq i \leq N)$ is equipped with a cognitive radio for data transmissions and a normal radio for control signals. We assume that the nodes are stationary and aware of their geographic positions. All transmitted data packets contain the information of their destinations' locations.

We assume that each SU adopts the spectrum sensing technology (e.g., feature detection [43]) to detect the channel availability in one location at one time, and it has the capability of channel switching at the packet level for data transmissions. It is also assumed that there exists a common control channel available for all nodes to exchange the control messages. According to [44] and [45], the temporal channel usage pattern of PTs over channel c_j can be characterized by an independent exponential ON/OFF state model with an activity probability $p_{c_j} = E[T_{on}^{c_j}]/(E[T_{on}^{c_j}] + E[T_{off}^{c_j}])$, where $E[T_{on}^{c_j}]$ is the average duration of the PU activity period, and $E[T_{off}^{c_j}]$ is the average duration of the PU silence period. Accordingly, the probability that PU does not transmit on channel c_j is $1 - p_{c_j}$. The mean duration of the channel spent in either ON/OFF state is assumed to follow an exponential distribution with the mean τ. Hence, the expected available transmission time is defined as $(1 - p_{c_j})\tau/p_{c_j}$. For each available channel, the SU keeps tracking of the elapsed time τ_{elapse}. Thus, the expected available transmission time can be calculated as $ATT(c_j) = (1 - p_{c_j})\tau/p_{c_j} - \tau_{elapse}$. This assumption is reasonable that the related work reported in [3] allows SUs to monitor and utilize past channel histories, to estimate the future spectrum availability.

The CRAHN is modeled by an undirected graph $G(V, E)$, where V is the set of CR nodes, and E is the set of links connecting the nodes. Let C_v denote the set of channels which are available to node $v \in V$. We adopt the protocol interference model and assume that all CR nodes have the same transmission and interference ranges. In CRAHNs, a link $l_{uv} \in E$ between node u and node v is considered to exist, provided that they have at least one common channel and the Euclidean distance between them is less than the transmission range. Besides, the communication links of our system are bidirectional, following the IEEE 802.11 model which uses link-layer ACKs to confirm delivery. Due to the unreliable nature of the wireless medium, each link in our system is associated with a packet delivery ratio (*PDR*), which is the ratio of the packets successfully delivered to a destination compared to the number of packets transmitted.

In this work, we only focus on the network layer without considering MAC layer issues such as synchronization, scheduling, collision avoidance and hidden terminal problem. In our method, we adopt the MAC protocol designed for opportunistic routing [6], [8], [46] to realize the coordination. Since we assume that the available channels are known to SUs, the scheduling problem of the MAC layer in our context is similar to that in ad hoc networks, which is a well-studied problem. Thus, it is reasonable to assume that there exist some idealistic scheduling mechanisms to reduce the overhead induced by relay contention and coordination at the MAC layer, which are out of the scope of this paper. To simplify the model, we have two more assumptions. One is that we assume perfect overhearing. That is, each node guarantees that the "unwanted" overheard packets are correctly received. The other is that we assume all nodes are always willing to cooperate.

B. PROBLEM STATEMENT

We aim to minimize the total number of transmissions for a given set of packets. The total number of transmissions to deliver a set of packets for a number of source-destination pairs is defined as the sum of the number of packets transmitted by all the nodes that participate in transmitting the set of packets. We let M be the number of packets to be transferred from the source to the destination and T_i be the number of packets each node i really delivered. Then, the average number of transmissions required to deliver one packet, denoted by $P_{average}$, can be written as

$$P_{average} = \frac{\sum_{i=1}^{N} T_i}{M} \tag{1}$$

Suppose that, given a set of SUs and available channels, our task is to design a routing scheme utilizing network coding and geographic opportunistic routing, such that $P_{average}$ in (1) is minimized.

IV. NETWORK CODING BASED GSOR SCHEME

In this section, we first provide a high-level overview of the network coding based GSOR scheme, and then describe its major components individually.

A. SCHEME OVERVIEW

In GSOR, a route from the source s to the destination d consists of multiple opportunistic route segments. As shown in Fig. 1, each segment has a tmp-source (temporary source), a tmp-destination (temporary destination) and a set of potential relay nodes. In GSOR, each potential relay node must have one-hop connection to the tmp-source and the tmp-destination (e.g., node $r_2^{T_0}$ in Fig. 1 is connected with node s and node m_1), which has essentially two reasons: 1) Due to the frequently changing nature of the primary frequency, the spectrum resources may be available for a short time. Thus, it is only possible to guarantee a sing-hop connection between two nodes for a limited duration. 2) In opportunistic routing, the source, the destination and the forwarding candidates should operate on the same channel for overhearing. If the forwarding candidates are not close to both the tmp-source and the tmp-destination, it is difficult to find one common available channel for implementing opportunistic routing, and it is impossible to guarantee that the tmp-destination can receive all packets successfully before the channel becomes unavailable.

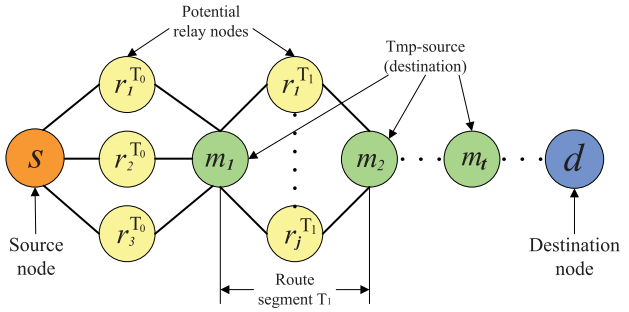


FIGURE 1. An overview of GSOR.

As illustrated in Fig. 1, GSOR works as follows: Initially, s is the tmp-source of the first route segment T_0. Then, the tmp-destination m_1 and its corresponding opportunistic relay set $\{r_1^{T_0}, r_2^{T_0}, r_3^{T_0}\}$ are chosen. After m_1 has obtained the full set of packets, it becomes the tmp-source of the next segment. Accordingly, the next tmp-destination m_2 and a new set of potential relays will be chosen to implement the opportunistic transmission procedure for T_1. The transmission processes over each segment continue until d receives all packets. In GSOR, the potential opportunistic relays for each segment are chosen before packet transmissions. However, according to the principle of opportunistic routing, the actual forwarding relay for transmitting each packet is not predetermined but opportunistically decided on the fly. Moreover, GSOR does not require the maintenance and establishment of the whole path from the source to the destination. The node which is closest to the final destination with desirable performance will be selected as the temp-destination.

In GSOR, multiple unicast flows may intersect at the same relay node, which brings network coding opportunities. An example of this case can be found in Fig. 2. In Fig. 2(a), one flow is from s_1 to d_1, and the other is from s_2 to d_2.

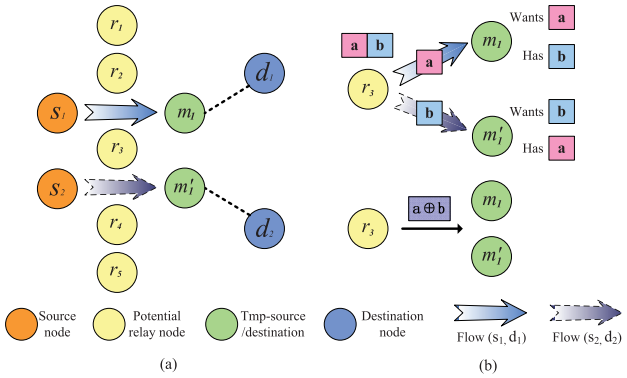


FIGURE 2. Illustration of network coding opportunities in GSOR.

We suppose that these two flows share one relay node r_3. As shown in Fig. 2(b), we consider the case that node r_3 has delivered two packets a and b to m_1 and m'_1 respectively, but neither m_1 or m'_1 successfully received their needed packets. Due to the spatial diversity and broadcast nature of wireless channels, multiple forwarders are likely to receive or overhear different packets at each moment. Thus, it is reasonable to suppose that node m_1 has the packet b and node m'_1 has the packet a, which may be obtained from overhearing or previous transmissions for other flows. In this case, instead of transmitting two packets to m_1 and m'_1 respectively, the relay node r_3 can broadcast one coded packet $a \oplus b$ to m_1 and m'_1, thus reducing one transmission compared with the case without network coding. In our scheme, by matching what each neighbor wants with what another neighbor has, each relay node can recognize coding opportunities and exploit them to forward multiple packets in a single transmission.

Briefly, the network coding based GSOR scheme is composed of three main components:

- **Forwarding set selection scheme:** This scheme describes how a tmp-source (source) node determines the optimal forwarding set from a subset of neighbors, and how to assign priorities to the forwarders in the set. We shall present this scheme in subsection IV-C.
- **Temporary destination selection scheme:** This scheme explains how the temp-destination is selected in each segment. The scheme exploits the location information instead of topology information to route packets to gradually approach and eventually reach the ultimate destination. In each segment, the scheme relies on a geographic greedy forwarding strategy to deliver packets to a locally optimal temp-destination. We shall present this scheme in subsection IV-D.
- **Network coding scheme:** This scheme describes how to mix as many packets as possible into a single coded transmission and discusses some practical implementation issues of combining network coding with GSOR. We shall present this scheme in subsection IV-E.

Before diving into specific routing schemes, it is crucial to define an appropriate routing metric for evaluating different opportunistic route segments. Moreover, it is essential to analyze the properties of the metric that can help us understand

how selection and prioritization of the forwarding candidate affect the performance. The theoretical analysis results enable us to design efficient routing algorithms that close to the optimal solution. In the next subsection IV-B, we shall introduce the routing metric and theoretical analysis results.

B. ROUTING METRIC AND PROPERTIES

1) THE SPDR METRIC

We define a new metric, called *SPDR*, to measure the packet delivery ratio of each opportunistic route segment, which is defined as below:

Definition 1: The Route Segment's Packet Delivery Ratio (SPDR) of an opportunistic route segment is the ratio of data packets received by the tmp-destination over the total data packets sent by the tmp-source.

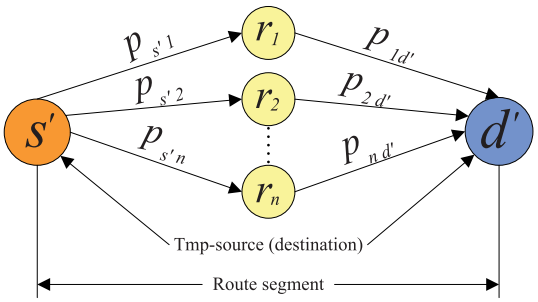


FIGURE 3. An example of one opportunistic route segment.

For clear presentation, we use an opportunistic route segment example in Fig. 3 to describe how to compute the *SPDR* metric. Let F denote the set of all nodes that have connections with both s' and d' in the same available channel at the same time instance. As shown in Fig. 3, an ordered set of forwarding candidates, denoted by $\pi(n) = \{r_1, \ldots, r_n\}$ with relay priority $r_1 > \ldots > r_n$, is selected from F. Then, the $SPDR(\pi(n))$ metric can be computed as follow:

$$SPDR(\pi(n)) = \sum_{i=1}^{n} p_{s'i} \cdot p_{id'} \cdot q_i \tag{2}$$

where the notation p_{uv} represents the packet delivery ratio of the link l_{uv}, and the notation q_i is the probability of none of the higher-priority nodes in $\pi(n)$ successfully receiving the packet from s' which is calculated as follow:

$$q_i = \begin{cases} \prod_{k=1}^{i-1}(1 - p_{s'k}), & i > 1 \\ 1, & i = 1 \end{cases} \tag{3}$$

2) PROPERTIES OF THE SPDR METRIC

In this part, we analyze and prove the properties of *SPDR*, which can help us to find the optimal solution for selecting the best forwarding relays for each opportunistic route segment. To simplify the notation and discussion, we provide the following definition.

Definition 2: Define $SPM(F, n)$ be the maximum SPDR achieved by selecting n forwarding candidates from F.

Then, we have the following theorems for $SPM(F, n)$.

Theorem 1: Given a forwarding candidate set $\pi(n)$, the $SPM(F, n)$ can only be achieved by assigning higher relay priorities to the forwarding nodes with larger values of $p_{id'}$ in (2).

Proof: In the following proofs, we assume that different nodes have different values of $p_{id'}$ (otherwise, there is no need to prioritize). The proof works by induction on n.

If $n = 1$, there is only one node involved, and then apparently Theorem 1 holds.

Next, we consider the case that $n = 2$, where two relay nodes r_x are r_y are involved. Assume $\pi^{xy}(2) = \{r_x, r_y\}$ and $\pi^{yx}(2) = \{r_y, r_x\}$ are two ordered node sets with different forwarding priorities $r_x > r_y$ and $r_y > r_x$, respectively. If $p_{xd'} > p_{yd'}$, we have

$$\begin{aligned} &SPDR(\pi^{xy}(2)) - SPDR(\pi^{yx}(2)) \\ &= p_{s'x} \cdot p_{xd'} + (1 - p_{s'x}) \cdot p_{s'y} \cdot p_{yd'} \\ &\quad - p_{s'y} \cdot p_{yd'} - (1 - p_{s'y}) \cdot p_{s'x} \cdot p_{xd'} \\ &= p_{s'x} \cdot p_{s'y} \cdot (p_{xd'} - p_{yd'}) > 0 \end{aligned} \tag{4}$$

Thus, Theorem 1 holds for $n = 2$.

Assume that Theorem 1 holds for $n = N (N \geq 2)$. By induction, we will prove it for $n = N + 1$. Suppose the relay nodes in $\pi^*(N + 1)$ with the permutation $\{r_1, r_2, \ldots, r_N, r_{N+1}\}$ can achieve $SPM(F, n)$ following $p_{1d'} > p_{2d'}, \ldots, p_{Nd'} > p_{(N+1)d'}$. If we assign one node $r_m \in \pi^*(N+1)$ with a different sequence number $t (1 \leq t \leq N+1, t \neq m)$, then the set $\pi'(N + 1)$ with the new permutation is built. To prove Theorem 1 holds for $n = N + 1$, we need to prove, for any new permutation, $SPDR(\pi^*(N + 1)) > SPDR(\pi'(N + 1))$ always holds. To prove this, we should distinguish two cases as follow.

Case 1: $1 \leq t < m$. This case indicates that we assign a lower sequence number (higher priority level) t for r_m. Clearly, node m_1 cannot be assigned with a lower sequence number than 1. Thus, in this case, we have $1 < m \leq N + 1$.

Case 2: $m < t \leq N + 1$. This case indicates that the node r_m is assigned with a higher sequence number (lower priority level) t. For the same reason as in Case 1, we also have $1 \leq m < N$ for Case 2.

Then, we shall prove that Theorem 1 holds for Case 1. Initially, we have

$$\begin{aligned} SPDR(\pi^*(N+1)) = \eta \cdot &\sum_{i=1}^{t-1} \cdot p_{s'i} \cdot p_{id'} \cdot q_i \\ &+ \theta \cdot \sum_{i=m+1}^{N+1} \cdot p_{s'i} \cdot p_{id'} \cdot q_i + q_t \cdot (p_{s't} \cdot p_{td'} \\ &+ \sum_{u=t+1}^{m-1} \cdot p_{s'u} \cdot p_{ud'} \cdot \prod_{i=t}^{u-1} (1 - p_{s'i}) \\ &+ p_{s'm} \cdot p_{md'} \cdot \prod_{i=t}^{m-1} (1 - p_{s'i})) \end{aligned} \tag{5}$$

and

$$SPDR(\pi'(N+1)) = \eta\cdot\sum_{i=1}^{t-1}\cdot p_{s'i}\cdot p_{id'}\cdot q_i + \theta\cdot\sum_{i=m+1}^{N+1}\cdot p_{s'i}\cdot p_{id'}\cdot q_i + q_t\cdot(p_{s'm}\cdot p_{md'} + (1-p_{s'm})\cdot(p_{s't}\cdot p_{td'} + \sum_{u=t+1}^{m-1}\cdot p_{s'u}\cdot p_{ud'}\cdot\prod_{i=t}^{u-1}(1-p_{s'i}))) \quad (6)$$

where if $t = 1$, then $\eta = 0$; if $m = N+1$, then $\theta = 0$; in other cases, η and θ are both equal to 1. Subtracting $SPDR(\pi'(N+1))$ from $SPDR(\pi^*(N+1))$, we have

$$\begin{aligned} &SPDR(\pi^*(N+1)) - SPDR(\pi'(N+1)) \\ &= q_t\cdot(p_{s'm}\cdot(p_{s't}\cdot p_{td'} + \sum_{u=t+1}^{m-1}\cdot p_{s'u}\cdot p_{ud'}\cdot\prod_{i=t}^{u-1}(1-p_{s'i}) \\ &\quad + p_{md'}\cdot\prod_{i=t}^{m-1}(1-p_{s'i})) - p_{s'm}\cdot p_{md'}) \end{aligned} \quad (7)$$

Since $p_{td'} > \ldots > p_{md'}$, then we have

$$\begin{aligned} &p_{s't}\cdot p_{td'} + \sum_{u=t+1}^{m-1}\cdot p_{s'u}\cdot p_{ud'} \\ &\quad\cdot\prod_{i=t}^{u-1}(1-p_{s'i}) + p_{md'}\cdot\prod_{i=t}^{m-1}(1-p_{s'i}) \\ &> p_{s't}\cdot p_{md'} + \sum_{u=t+1}^{m-1}\cdot p_{s'u}\cdot p_{md'}\cdot\prod_{i=t}^{u-1}(1-p_{s'i}) \\ &\quad + p_{md'}\cdot\prod_{i=t}^{m-1}(1-p_{s'i}) \\ &= p_{md'}\cdot(p_{s't} + \sum_{u=t+1}^{m-1}\cdot p_{s'u}\cdot\prod_{i=t}^{u-1}(1-p_{s'i}) + \prod_{i=t}^{m-1}(1-p_{s'i})) \\ &\quad\vdots \\ &= p_{md'}\cdot(p_{s't} + p_{s'(t+1)}\cdot(1-p_{s't}) + (1-p_{s'(t+1)})\cdot(1-p_{s't}) \\ &= p_{md'}\cdot(p_{s't} + (1-p_{s't})) = p_{md'} \end{aligned} \quad (8)$$

Finally, combining (24) and (8), we have

$$\begin{aligned} &SPDR(\pi^*(N+1)) - SPDR(\pi'(N+1)) \\ &\quad > q_t\cdot(p_{s'm}\cdot p_{md'} - p_{s'm}\cdot p_{md'}) = 0 \end{aligned} \quad (9)$$

Now we have completed the proof for Case 1. We can use the same procedure above to prove it for Case 2, which is the symmetry of Case 1. Consequently, Theorem 1 holds for $n = N+1$ and this completes the induction. ■

Applying Theorem 1, the following theorem is obtained.

Theorem 2: The ordered set $\pi^*(n)$ that achieves SPM (F, n) is a subset of the sequenced set $\pi^*(n+1)$ that achieves $SPM(F, n+1)$.

Proof: Firstly, we assume that $\Re = \{r_1, \ldots, r_m\}$ is the corresponding well prioritized set of F based on $p_{id'}$, where the subscript of r stands for the priority order. According to Theorem 1, no matter how many nodes are involved in $\pi(n)$, the relay nodes in $\pi(n)$ are arranged with the same permutation as in $\Re$.

We then prove $\pi^*(n) \subset \pi^*(n+1)$ $(1 \le n < |F|)$ by mathematical induction on n.

If n=1, assuming $\pi^*(1) = \{r_a\}$, then we have

$$\begin{aligned} SPM(F, 1) &= SPDR(\pi^*(1)) \\ &= p_{s'a}\cdot p_{ad'} = \max\{p_{s'i}\cdot p_{id'}\}(\forall r_i \in \Re) \end{aligned} \quad (10)$$

If n=2, we prove it by contradiction. Suppose $\pi^*(2) = \{r_b, r_c\}$ can achieve the maximum value $SPM(F, 2)$, and $\pi^*(1) = \{r_a\} \not\subset \pi^*(2)$. Then, we have

$$\begin{aligned} SPM(F, 2) &= SPDR(\pi^*(2)) \\ &= p_{s'b}\cdot p_{bd'} + (1-(p_{s'b}))\cdot p_{s'c}\cdot p_{cd'} \\ &= \max\{p_{s'i}\cdot p_{id'} + (1-p_{s'i})\cdot p_{s'j}\cdot p_{jd'}\} \\ &\quad (\forall r_i, r_j \in \Re, i < j) \end{aligned} \quad (11)$$

We first consider the case that $a<b$, which indicates that the priority of r_a is higher than r_b. Let r_a replace r_c to construct a new set $\pi^{a,b}(2) = \{r_a, r_b\}$, according to (11), we have

$$\begin{aligned} &SPDR(\pi^{a,b}(2)) - SPDR(\pi^*(2)) \\ &= p_{s'a}\cdot p_{ad'} + (1-p_{s'a})\cdot p_{s'b}\cdot p_{bd'} - p_{s'b}\cdot p_{bd'} \\ &\quad - (1-p_{s'b})\cdot p_{s'c}\cdot p_{cd'} \\ &= p_{s'a}\cdot p_{ad'} - p_{s'a}\cdot p_{s'b}\cdot p_{bd'} - (1-p_{s'b})\cdot p_{s'c}\cdot p_{cd'} \end{aligned} \quad (12)$$

According to (10), it implies that $p_{s'a}\cdot p_{ad'} > p_{s'c}\cdot p_{cd'}$. Moreover, notice that the priority order is $r_a > r_b > r_a$, which means $p_{ad'} > p_{bd'} > p_{cd'}$, then we have

$$\begin{aligned} &p_{s'a}\cdot p_{ad'} - p_{s'a}\cdot p_{s'b}\cdot p_{bd'} - (1-p_{s'b})\cdot p_{s'c}\cdot p_{cd'} \\ &\quad > p_{s'a}\cdot p_{ad'} - p_{s'a}\cdot p_{s'b}\cdot p_{ad'} - (1-p_{s'b})\cdot p_{s'c}\cdot p_{cd'} \\ &\quad > p_{s'a}\cdot p_{ad'}\cdot(1-p_{s'b}) - p_{s'c}\cdot p_{cd'}\cdot(1-p_{s'b}) \\ &\quad > (1-p_{s'b})\cdot(p_{s'a}\cdot p_{ad'} - p_{s'c}\cdot p_{cd'}) > 0 \end{aligned} \quad (13)$$

Based on (12) and (13), we can conclude that $SPDR(\pi^{a,b}(2)) > SPDR(\pi^*(2))$, which contradicts with the assumption that $\pi^*(2) = \{r_b, r_c\}$ can achieve the maximum.

Then, we consider the case that $a>b$. Similarly, we can replace r_c with r_a to construct a new set $\pi^{b,a}(2) = \{r_b, r_a\}$. According to (10), we always have $SPDR(\pi^{b,a}(2)) > SPDR(\pi^*(2))$, which also contradicts with the assumption. Thus, Theorem 2 holds for $n = 2$.

Suppose that the theorem holds for $n = N(N \ge 2)$. Then, we shall prove that it also holds for $n = N+1$, which means $\pi^*(N) \subset \pi^*(N+1)$. To facilitate the following proof, we assume that $\pi^*(N) = \{r_1, \ldots, r_N\}$ is one permutation of nodes with relay priority $r_1 > \ldots > r_N$, which can produce the maximum. In addition, we assume that $\pi^*(N+1) = \{\hat{r}_1, \ldots, \hat{r}_{N+1}\}$ is also a set of ordered nodes with relay priority $\hat{r}_1 > \ldots > \hat{r}_{N+1}$, which can achieve the maximum. To prove this, we consider two cases as follow:

Case 1: $\hat{r}_1 \notin \pi^*(N)$, which means that the first node in $\pi(N+1)$ cannot be any node in $\pi(N)$. Since $SPDR(\pi^*(N))$ can achieve the maximum; it is undoubtedly that $\{\hat{r}_1, \pi^*(N)\}$ should be $\pi^*(N+1)$. Thus, $\pi^*(N) \subset \pi^*(N+1)$ and the theorem holds for $n = N+1$.

Case 2: $\hat{r}_1 \in \pi^*(N)$, which means that the first node in $\pi(N+1)$ is selected from $\pi(N)$. To prove it, we first prove $\hat{r}_1 = r_1$ by contradiction. In the following, to distinguish PDR values for $r_i \in \pi(N)$ and $\hat{r}_j \in \pi(N+1)$, we use the notations $p_{s'r_i(\hat{r}_j)}$ and $p_{r_i(\hat{r}_j)d'}$ instead of $p_{s'i}$ and $p_{id'}$.

Firstly, we suppose $\hat{r}_1 = r_2$, which indicates that $\pi^*(N+1) = \{r_2, \hat{r}_2, \ldots, \hat{r}_{N+1}\}$ can produce the maximum. In this situation, the node r_1 cannot be included in $\pi^*(N+1)$, and at least two additional nodes should be added into $\pi^*(N+1)$, which are not included in $\pi^*(N)$. Assume the first and the second node in $\pi^*(N+1)$ but not in $\pi^*(N)$ are $\hat{r}_j$ and $\hat{r}_k$ respectively. For convenient description, we consider the general case that $\pi^*(N+1) = \{r_2, \ldots, r_j, \hat{r}_j, r_{j+1} \ldots, r_{k-1}, \hat{r}_k, \hat{r}_{k+1} \ldots, \hat{r}_{N+1}\}$. Here, we use $\mathcal{A}$ to denote the ordered set $\{r_2, \ldots, r_j, \hat{r}_j, r_{j+1} \ldots, r_{k-1}\}$, and use $\mathcal{B}$ to denote the set $\{\hat{r}_k, \hat{r}_{k+1} \ldots, \hat{r}_{N+1}\}$, then we have

$$\begin{aligned} SPDR(\pi^*(N+1)) = {} & SPDR(\mathcal{A}) \\ & + \prod_{i=2}^{j}(1 - p_{s'r_i})\cdot(1 - p_{s'\hat{r}_j}) \\ & \cdot \prod_{q=j+1}^{k-1}(1 - p_{s'r_q})\cdot SPDR(\mathcal{B}) \end{aligned} \tag{14}$$

Then, we construct a new ordered set $\pi'(N+1) = \{r_1, \ldots, r_j, r_{j+1} \ldots, r_{k-1}, \hat{r}_k, \hat{r}_{k+1} \ldots, \hat{r}_{N+1}\}$. Let $\mathcal{C}$ denote the ordered set $\{r_1, \ldots, r_j, r_{j+1} \ldots, r_{k-1}\}$, then we have

$$\begin{aligned} SPDR(\pi'(N+1)) = {} & SPDR(\mathcal{C}) \\ & + \prod_{i=1}^{k-1}(1 - p_{s'r_i})\cdot SPDR(\mathcal{B}) \end{aligned} \tag{15}$$

Let $\mathcal{D}$ denote an ordered set $\{r_k, r_{k+1} \ldots, r_N\}$, and we suppose $\pi'(N) = \{\mathcal{A}, \mathcal{D}\}$. Since $SPDR(\pi^*(N+1)) > SPDR(\pi'(N+1))$, then we have the following inequality by replacing $\mathcal{B}$ with $\mathcal{D}$ in (14) and (15).

$$\begin{aligned} & SPDR(\pi'(N)) \\ & = SPDR(\mathcal{A}) + \prod_{i=2}^{j}(1 - p_{s'r_i})\cdot(1 - p_{s'\hat{r}_j}) \\ & \quad \cdot \prod_{q=j+1}^{k-1}(1 - p_{s'r_q})\cdot SPDR(\mathcal{D}) \\ & > SPDR(\mathcal{C}) + \prod_{i=1}^{k-1}(1 - p_{s'r_i})\cdot SPDR(\mathcal{D}) \\ & = SPDR(\pi^*(N)) \end{aligned} \tag{16}$$

The inequality (16) contradicts with the fact that the $SPDR(\pi^*(N))$ is the maximum. Thus, the assumption that $\hat{r}_1 = r_2$ is wrong. Then, by the same approach, we can also prove that $\hat{r}_1 \notin \{r_3, \ldots, r_N\}$. Accordingly, we can conclude that $r_1 \in \pi^*(N+1)$ and $\hat{r}_1 = r_1$.

By the inductive hypothesis, we can also consider those two cases for $\hat{r}_2, \ldots \hat{r}_{N+1}$ and conclude that $\pi^*(N) - \{r_1\} \subset \{\hat{r}_2, \ldots \hat{r}_{N+1}\}$. Therefore, it is true that the theorem holds for $n = N+1$. ■

Applying Theorem 1 and Theorem 2, we have the following corollaries.

Corollary 1: $SPM(F, n)$ is a strictly increasing function of n. That is, $SPM(F, n+1) - SPM(F, n) > 0$, $\forall n$, $1 \le n < |F|$.

Proof: Suppose we have $\pi^*(n) = \{r_1, \ldots, r_n\}$ and $\pi^*(n+1) = \{\hat{r}_1, \ldots, \hat{r}_{n+1}\}$. According to Theorem 2, it is reasonable to assume that $\pi^*(n+1) - \pi^*(n) = \hat{r}_x$. Applying Theorem 1, we have

$$\begin{aligned} SPDR(\pi^*(n)) & < SPDR(\pi^*(n)) + \prod_{i=1}^{n}(1 - p_{s'r_i})\cdot p_{s'\hat{r}_j}\cdot p_{\hat{r}_j d'} \\ & = SPDR(\{r_1, \ldots, r_n, \hat{r}_j\}) \le SPDR(\pi^*(n+1)) \end{aligned} \tag{17}$$

■

Corollary 2: $SPM(F, n)$ is a concave function of n. That is, $SPM(F, n+1) - SPM(F, n) < SPM(F, n) - SPM(F, n-1)$, $\forall n$, $2 \le n < |F|$.

Proof: Based on Theorem 2, we can assume that $\pi^*(n) - \pi^*(n-1) = r_x$ and $\pi^*(n+1) - \pi^*(n) = \hat{r}_y$. Then we prove the corollary by two cases.

Case 1: $p_{r_x d'} > p_{\hat{r}_y d'}$. In this case, the three ordered sets can be represented as $\pi^*(n+1) = \{\mathcal{A}, r_x, \mathcal{B}, \hat{r}_y, \mathcal{C}\}$, $\pi^*(n) = \{\mathcal{A}, r_x, \mathcal{B}, \mathcal{C}\}$, and $\pi^*(n-1) = \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$. For ease of presentation, we consider the general case that $\mathcal{A} = \{r_1, \ldots r_j\}$, $\mathcal{B} = \{r_{j+1}, \ldots r_k\}$, and $\mathcal{C} = \{r_{k+1}, \ldots r_N\}$. For convenient description, we use $\overline{\mathcal{A}}$ to denote $\prod_{q=1}^{j}(1 - p_{s'r_q})$. Since $SPDR(\pi^*(n)) > SPDR(\{\mathcal{A}, \mathcal{B}, \hat{r}_y, \mathcal{C}\})$, then we have

$$\begin{aligned} & SPDR(\pi^*(n+1)) - SPDR(\pi^*(n)) \\ & \quad < SPDR(\{\mathcal{A}, r_x, \mathcal{B}, \hat{r}_y, \mathcal{C}\}) - SPDR(\{\mathcal{A}, \mathcal{B}, \hat{r}_y, \mathcal{C}\}) \\ & \quad = \overline{\mathcal{A}}\cdot p_{s'r_x}\cdot p_{r_x d'} + \overline{\mathcal{A}}\cdot(1 - p_{s'r_x})\cdot SPDR(\{\mathcal{B}, \hat{r}_y, \mathcal{C}\}) \\ & \qquad - \overline{\mathcal{A}}\cdot SPDR(\{\mathcal{B}, \hat{r}_y, \mathcal{C}\}) \\ & \quad = \overline{\mathcal{A}}\cdot p_{s'r_x}\cdot(p_{r_x d'} - SPDR(\{\mathcal{B}, \hat{r}_y, \mathcal{C}\})) \end{aligned} \tag{18}$$

and

$$\begin{aligned} & SPDR(\pi^*(n)) - SPDR(\pi^*(n-1)) \\ & \quad = \overline{\mathcal{A}}\cdot p_{s'r_x}\cdot p_{r_x d'} + \overline{\mathcal{A}}\cdot(1 - p_{s'r_x})\cdot SPDR(\{\mathcal{B}, \mathcal{C}\}) \\ & \qquad - \overline{\mathcal{A}}\cdot SPDR(\{\mathcal{B}, \mathcal{C}\}) \\ & \quad = \overline{\mathcal{A}}\cdot p_{s'r_x}\cdot(p_{r_x d'} - SPDR(\{\mathcal{B}, \mathcal{C}\})) \end{aligned} \tag{19}$$

Corollary 1 shows that $SPDR(\{\mathcal{B}, \hat{r}_y, \mathcal{C}\})$ is larger than $SPDR(\{\mathcal{B}, \mathcal{C}\})$. Thus, from (18) and (19), corollary 2 holds.

Case 2: $p_{\hat{r}_y d'} > p_{r_x d'}$. Similarly, the sets can be represented as $\pi^*(n+1) = \{\mathcal{A}, \hat{r}_y, \mathcal{B}, r_x, \mathcal{C}\}$, $\pi^*(n) = \{\mathcal{A}, \mathcal{B}, r_x, \mathcal{C}\}$, and $\pi^*(n-1) = \{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$. We adopt the same definitions of $\mathcal{A}(\overline{\mathcal{A}})$, $\mathcal{B}$, and $\mathcal{C}$ as those in Case 1. Notice that

$SPDR(\{\mathcal{A}, \hat{r}_y, \mathcal{B}, \mathcal{C}\}) < SPDR(\pi^*(n))$, then we have

$$\begin{aligned} &SPDR(\pi^*(n)) - SPDR(\pi^*(n-1)) \\ &> SPDR(\{\mathcal{A}, \hat{r}_y, \mathcal{B}, \mathcal{C}\}) - SPDR(\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}) \\ &= \overline{\mathcal{A}} \cdot p_{s'\hat{r}_y} \cdot p_{\hat{r}_y d'} + \overline{\mathcal{A}} \cdot (1 - p_{s'\hat{r}_y}) \cdot SPDR(\{\mathcal{B}, \mathcal{C}\}) \\ &\quad - \overline{\mathcal{A}} \cdot SPDR(\{\mathcal{B}, \mathcal{C}\}) \\ &= \overline{\mathcal{A}} \cdot p_{s'\hat{r}_y} \cdot (p_{\hat{r}_y d'} - SPDR(\{\mathcal{B}, \mathcal{C}\})) \end{aligned} \tag{20}$$

and

$$\begin{aligned} &SPDR(\pi^*(n+1)) - SPDR(\pi^*(n)) \\ &= \overline{\mathcal{A}} \cdot p_{s'\hat{r}_y} \cdot p_{\hat{r}_y d'} + \overline{\mathcal{A}} \cdot (1 - p_{s'\hat{r}_y}) \cdot SPDR(\{\mathcal{B}, r_x, \mathcal{C}\}) \\ &\quad - overline\mathcal{A} \cdot SPDR(\{\mathcal{B}, r_x, \mathcal{C}\}) \\ &= \overline{\mathcal{A}} \cdot p_{s'\hat{r}_y} \cdot (p_{\hat{r}_y d'} - SPDR(\{\mathcal{B}, r_x, \mathcal{C}\})) \end{aligned} \tag{21}$$

Corollary 1 shows that $SPDR(\{\mathcal{B}, r_x, \mathcal{C}\}) > SPDR(\{\mathcal{B}, \mathcal{C}\})$. Then, combing inequality (20) with (21), we can conclude that Corollary 2 holds. ■

C. FORWARDING SET SELECTION SCHEME

In this subsection, we first derive the lower bound of *SPDR*. Then, we propose an optimal forwarding set selection algorithm to reach the upper bound of *SPDR*, and propose a sub-optimal solution with less computation complexity.

1) THE LOWER BOUND OF SPDR

The lower bound of *SPDR* is described as follow.

Proposition 1: The lower bound of $SPDR(\pi(n))$ is a special case that includes one node r^* in the forwarding set and given by (22).

$$SPDR_{lower} = p_{s'r^*} \cdot p_{r^*d'} = \min\{p_{s'i} \cdot p_{id'}\} (\forall r_i \in F) \tag{22}$$

Proof: If $n = 1$, it is obvious that Proposition 1 holds.

If $n > 1$, we assume that $\pi(n) = \{r_1, \ldots, r_n\}$ is an ordered subset of F with priority $r_1 > \ldots > r_n$. Then, we consider two cases in the following:

Case 1: Node r^* cannot be the highest priority node in $\pi(n)$. Then, according to (22), we have

$$\begin{aligned} SPDR(\pi(n)) &= p_{s'1} \cdot p_{1d'} + \sum_{i=2}^{n} \cdot p_{s'i} \cdot p_{id'} \cdot q_i \\ &> p_{s'1} \cdot p_{1d'} \geq p_{s'r^*} \cdot p_{r^*d'} = SPDR_{lower} \end{aligned} \tag{23}$$

Case 2: Node r^* is the highest priority node in $\pi(n)$. In this case, we have

$$\begin{aligned} SPDR(\pi(n)) &= p_{s'r^*} \cdot p_{r^*d'} + \sum_{i=2}^{n} \cdot p_{s'i} \cdot p_{id'} \cdot q_i \\ &> p_{s'r^*} \cdot p_{r^*d'} = SPDR_{lower} \end{aligned} \tag{24}$$

Thus, Proposition 1 always holds. ■

2) THE UPPER BOUND OF SPDR AND THE OPTIMAL FORWARDING SET SELECTION ALGORITHM

To maximize $SPDR(\pi(n))$ and compute the upper bound, we can try each possible combinations of different subsets of the relay candidate set with different channels. However, if there are m available channels and n candidate relays, the exhaustive search running time is $O(m \cdot n! \cdot e)$.

To reduce the complexity, we design a more efficient solution to select the optimal forwarding set for each opportunistic route segment, which is called the Optimal Forwarding Set Selection (Optimal-OFSS) algorithm. As described in Algorithm 1, each time it includes one best node r^* with the largest value of *SPDR* into F^* (lines 3-9). More specifically, the nodes in the corresponding new set $\pi(n)$ are prioritized according to Theorem 1 (line 4), and the *SPDR* value is calculated by (2) (line 5). Then, based on the properties of $SPDR(\pi(n))$, we have the following proposition.

Algorithm 1 Optimal-OFSS

tmp-source

Input: s'; tmp-destination d'; available node set F.

Output: The optimal forwarding set F^* and SPM^* for the opportunistic route segment from s' to d'.

1: $F^* \leftarrow \emptyset$; $\pi(n) \leftarrow \emptyset$; $r^* \leftarrow 0$; $SPM^* \leftarrow 0$;
2: **for** $n = 1$ to $|F|$ **do**
3: **for all** node $r_j \in F$ **do**
4: Include r_j and F^* as a new set $\pi(n)$, where the nodes in $\pi(n)$ are prioritized based on Theorem 1;
5: Calculate $SPDR(\pi(n))$ according to (2);
6: **if** $(SPDR(\pi(n)) > SPM^*)$ **then**
7: $SPM^* \leftarrow SPDR(\pi(n))$; $r^* \leftarrow r_j$;
8: **end if**
9: **end for**
10: $F^* \leftarrow F^* \cup r^*$; $F \leftarrow F \setminus r^*$;
11: **end for**
12: **return** (F^*, SPM^*)

Proposition 2: The optimal value SPM^* obtained in Algorithm 1 is an upper bound on SPDR.

Proof: In Algorithm 1, at each iteration, it inserts a new node r_j into the optimal forwarding set F^* containing $n-1$ nodes. Theorem 2 can prove that this greedy-adding heuristic can find an optimal forwarding set. In addition, based on Theorem 1, the priority rule can ensure that the maximum value of *SPDR* can be achieved for each iteration. Furthermore, according to Corollary 1, $SPM(F, n)$ is a strictly increasing function of n. Thus, the optimal value SPM^* can only be achieved by including all the nodes in F. ■

3) THE SUB-OPTIMAL FORWARDING SET SELECTION ALGORITHM

Nevertheless, Algorithm 1 is still not the best solution. Intuitively, the coordination overheads and computation complexities increase when more forwarding candidates are involved. Moreover, based on Corollary 2, although *SPM* keeps increasing by involving more relay nodes for cooperation, the gained extra progress becomes negligible. Hence, the disadvantage of increased overhead may overwhelm the *SPM* gain. To tackle this issue, we design a Sub-Optimal Forwarding Set Selection (Sub-Optimal-OFSS) algorithm to

approach the optimal, which is described in Algorithm 2. The difference between Algorithm 2 and Algorithm 1 is that Algorithm 2 compares the SPM^* obtained in the current loop with the SPM^- produced in the previous loop. If the gap between the SPM^* and the SPM^- is less than a threshold α, then we prevent adding any new node into F^* (line 3). Suppose that the size of F is n, then the complexity of the Sub-Optimal-OFSS algorithm is $O(n^2)$, which is much lower than the exhaustive search method.

Algorithm 2 Sub-Optimal-OFSS

tmp-source

Input: s'; tmp-destination d'; available node set F; threshold value α.

sub-optimal

Output: forwarding set F^* and SPM^* for the opportunistic route segment from s' to d'.

1: $F^* \leftarrow \emptyset$; $\pi(n) \leftarrow \emptyset$; $r^* \leftarrow 0$; $SPM^* \leftarrow 0$; $SPM^- \leftarrow -1$;
2: **for** $n = 1$ to $|F|$ **do**
3: **while** $(SPM^* - SPM^- \geq \alpha)$ **do**
4: $SPM^- \leftarrow SPM^*$;
5: Run lines 3-10 in Algorithm 1;
6: **end while**
7: **end for**
8: **return** (F^*, SPM^*)

Then, we proceed to prove that the solution obtained by Algorithm 2 is sub-optimal.

Proposition 3: The sub-optimal solution of Algorithm 2 is close to the optimal solution of Algorithm 1, and the gap between them is upper bounded by $(|F| - k) * \alpha$*, where* α *is the threshold and* k *is the number of nodes in the sub-optimal forwarding set.*

Proof: According to Algorithm 2 and Corollary 2, we have

$$SPM(F, n) - SPM(F, n-1) < \alpha \ (k \leq n \leq |F|) \quad (25)$$

Then, the gap between the sub-optimal value $SPM(F, k)$ and the optimal value $SPM(F, |F|)$ is expressed as follow:

$$\begin{aligned} &SPM(F, |F|) - SPM(F, k) \\ &\quad = SPM(F, |F|) - SPM(F, |F| - 1) \\ &\qquad + SPM(F, |F| - 1) - SPM(F, |F| - 2) \\ &\qquad \ldots + SPM(F, k + 1) - SPM(F, k) \\ &\quad \leq (|F| - k) * \alpha \end{aligned} \quad (26)$$

In Algorithm 2, with a smaller value of the threshold α, it can achieve a larger number k. Thus, when α becomes smaller, the gap value of $(|F| - k) * \alpha$ also becomes smaller and approaches zero. Then, it is concluded that the sub-optimal solution is close to the optimal solution. ■

D. TEMPORARY DESTINATION SELECTION SCHEME

In this subsection, we first express the trade-off between the *SPDR* and the distance advancement. Then, we introduce a local metric and propose a heuristic algorithm for selecting the temp-destination in each opportunistic route segment.

1) TRADE-OFF BETWEEN THE SPDR AND THE DISTANCE ADVANCEMENT

As shown in Fig. 1, starting from s, we need to choose a sequence of tmp-destinations (segments) leading to d. Unfortunately, this choice becomes challenging due to the trade-off between the *SPDR* and the distance advancement. On the one hand, we prefer the tmp-destination has the largest distance advancement towards d, thus it would take fewer segments to reach d. However, the distance between the tmp-source and the tmp-destination will be long in this case, and there will be very few eligible relay nodes in between them, which leads the *SPDR* very small. On the other hand, if the tmp-destination is selected too close to the tmp-source, it would take more segments to reach d, which makes little segment's distance advancement.

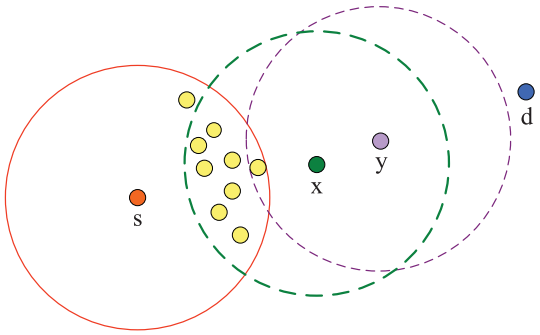


FIGURE 4. An example of greedy forwarding.

An example of this situation is shown in Fig. 4. If we choose y as the tmp-destination that has the largest advancement to d, y will be quite far away from s, which makes the relay node set small. Recall that the relay nodes between s and the tmp-destination must be within the transmission range of both s and the tmp-destination (i.e., the nodes falling into the overlapped region of two disks centered at s and the tmp-destination respectively). In this case, there is only one node falling into the transmission range of s and y. However, if we choose x as the tmp-destination, there are more nodes in the relay set.

2) HEURISTIC TEMPORARY DESTINATION SELECTION ALGORITHM

To balance the trade-off between the *SPDR* and the distance advancement, we introduce a new local metric, called packet delivery advancement (*PDA*), to evaluate each candidate tmp-destination u_k, which is calculated as follow:

$$PDA(m_t, u_k) = SPDR(m_t, u_k) * D(m_t, u_k) \quad (27)$$

where m_t is the tmp-source (as shown in Fig. 1) and $D(m_t, u_k)$ represents the distance advancement, which is defined as follow:

$$D(m_t, u_k) = d(m_t, d) - d(u_k, d) \quad (28)$$

where $d(m_t, d)$ and $d(u_k, d)$ are the Euclidean distances between m_t and d, and u_k and d, respectively.

In subsection IV-A, we have stated that each opportunistic relay node should have one-hop connection to the tmp-source and the tmp-destination. Thus, the tmp-destination should be selected from the two-hops neighboring set of the tmp-source m_t towards the final destination, which is denoted by N_{m_t}. Then, we propose the temporary destination selection algorithm, as shown in Algorithm 3. First, line 4 executes Algorithm 1 (or Algorithm 2) presented in subsection IV-C to determine the optimal (or sub-optimal) forwarding set of each $u_k \in N_{m_t}$ on each channel $c_j \in C$. Second, line 5 computes the *SPDR* of the forwarding set, and line 6 computes the *PDA* of the opportunistic route segment from m_t to each u_k. If the node u_k with channel c_j can produce the highest value of *PDA*, then the node u_k will be selected as the tmp-destination, and the channel c_j will be chosen as the operating channel (lines 7-8). Moreover, line 7 ensures that the expected available transmission time of the channel c^* must be adequate for each transmission, which means $ATT(c^*) \geq \beta$. Specifically, the *ATT* metric has been defined in subsection III-A, and the threshold β is decided by the packet size, the data rate and the number of packets of each batch. If the number of nodes in N_{m_t} is n, and the number of available channels in C is m, then the complexity of Algorithm 3 is $O(m{\cdot}n)$.

Algorithm 3 Temporary Destination Selection Algorithm

temp-source

Input: m_t; neighboring set N_{m_t}; available channel set C; threshold value β.

tmp-destination

Output: u_k^*; operation channel c^*;

$u_k^* \leftarrow 0$; $c^* \leftarrow 0$; $PDA^* \leftarrow 0$

2: **for all** $u_k \in N_{m_t}$ **do**

  for all $c_j \in C$ **do**

4: Run Algorithm 1 (or Algorithm 2) to find the optimal (or sub-optimal) forwarding set;

Compute the *SPDR* according to (2);

6: Compute the *PDA* according to (27);

if $(PDA > PDA^*)$ **and** $(ATT(c^*) \geq \beta)$ **then**

8: $u_k^* \leftarrow u_k$; $c^* \leftarrow c_j$

end if

10: **end for**

end for

12: **return** u_k^*

E. NETWORK CODING SCHEME

In the GSOR scheme, a relay node may relay the traffic for multiple flows. We consider the scenario that the relay node has a set of k packets $P = \{p_1, p_2, \ldots, p_k\}$ to be delivered to m tmp-destination nodes belonging to $D = \{d_1, d_2, \ldots, d_m\}$. Let $P(d_i)$ denote the set of packets that are destined to d_i, and let $H(d_i)$ denote the set of packets that d_i has overheard or received. Then, the relay node can XOR $t (t \leq k)$ packets together, on the condition that each d_i can decode its required packets by exploiting the packets in $H(d_i)$. Fig. 5 shows an example of this case, where each destination node has a set of packets it has overheard in its packet pool. Suppose that the relay node has knowledge of what the neighbors have overheard, then it can broadcast a coded packet $p_3 \oplus p_4 \oplus p_5 \oplus p_6$ to the destination nodes d_1, d_2, d_3 and d_4, who can decode their needed packet p_5, p_6, p_3 and p_4 respectively by using the packets in their own packet pool.

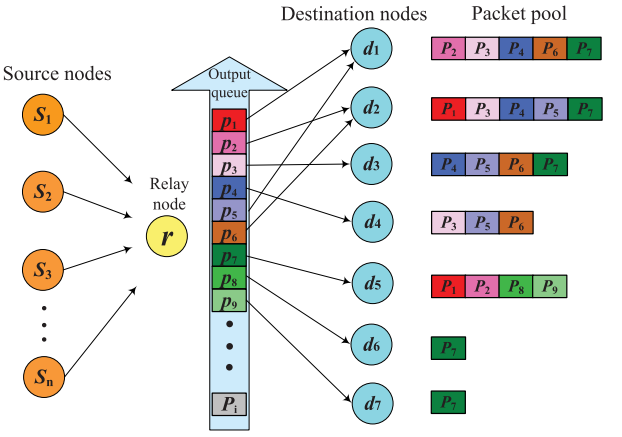


FIGURE 5. An example of network coding.

In this work, our task is to code the maximum number of packets in each transmission. In the following parts, we first construct a coding graph and show it is a reduction from the maximum clique problem. Then, we propose a heuristic coding scheme to achieve this goal.

1) CODING GRAPH CONSTRUCTION

In our scheme, each relay node r needs to decide which packets should be coded together that can reach the maximum benefit of coding. Here, we introduce a coding graph $G_p(V_p, E_p)$ to help make this decision. As for G_p, the vertex set represents the packets that r currently stored, and there is an edge $(p_i, p_j) \in E_p$ exists iff p_i and p_j can be coded together for one transmission. More specifically, considering two packets p_i and p_j are transmitted to d_i and d_j respectively, if $p_i \in H(d_j)$ and $p_j \in H(d_i)$, then p_i and p_j can be coded together. The corresponding coding graph of Fig. 5 is constructed as shown in Fig. 6. For example, in the figure, p_5 and p_6 are connected because the coded packet $p_5 \oplus p_6$ can be decoded by d_1 and d_2 successfully.

In the graph theory, a clique C is a subset of V_p, which is a complete subgraph of G_p and has the property that every two vertices in the subgraph are connected. Clearly, when the vertices belong to the same clique C, the corresponding packets can be coded (XORed) together and decoded at the tmp-destination. As mentioned above, our task is to code the

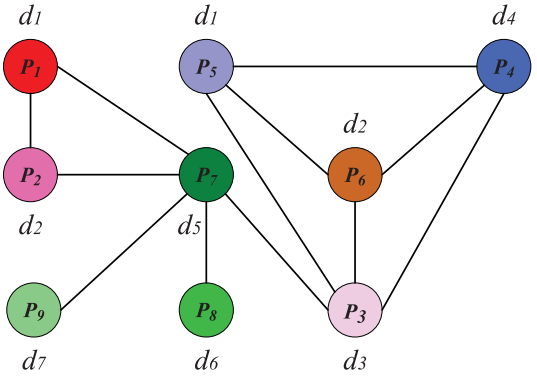


FIGURE 6. The corresponding coding graph of Fig. 5.

maximum number of packets in each transmission, which is equivalent to find the clique with the maximum number of vertices (packets). Thus, this coding problem can be converted into the maximum clique problem, which is proved to be NP-hard.

2) CODING STRATEGY

To solve the problem efficiently, we adopt a heuristic greedy approach to construct the maximal clique by repeatedly adding a vertex with the largest degree to a partial clique. In particular, let $d(v_i)$ denote the degree of the vertex v_i in G_p. We first choose the vertex v_i with the highest degree $d(v_i)$ and add it into an empty clique C. Then, we construct a subgraph $G'_p \subseteq G_p$, which only includes the vertices connected to v_i. After that, we select the next highest degree vertex from the remaining vertices except for v_i. If $C \cup \{v'_i\}$ is still a clique, we add v'_i into C. We repeat this selecting and adding process until no more vertex can be included.

Indeed, the aforementioned greedy method is efficient. However, the highest degree vertex v_i in G_p may not be a good start point, as indicated in the following situation. Actually, after generating the sub-graph G'_p including the first added vertex, the remaining vertices in G'_p may have a relatively low degree, which cannot be added to C further. Fig. 6 shows an example of this case. In the figure, the degree of vertex p_7 is 5 while vertex p_6 is 3. Adopting the coding method mentioned above, p_7 will be chosen as the start point. Unfortunately, the maximal clique in the graph is $\{p_3, p_4, p_5, p_6\}$, excluding p_7. To tackle this problem, we define another indicator as below, which is called effective degree $d^*(v_i)$, to select the vertices.

Definition 3: Suppose G'_p is a subgraph of G_p that only includes the vertices connecting v_i. The effective degree $d^*(v_i)$ is defined as the total number of vertices whose degrees in G'_p are above a threshold γ.

The following example is presented to illustrate this definition. As shown in Fig. 6, the effective degree of p_7 and p_6 are 2 and 3 respectively. If we set γ to be 2, then vertex p_6 will be firstly added to the clique and the maximal clique $\{p_3, p_4, p_5, p_6\}$ can be obtained.

3) IMPLEMENTATION ISSUES

In this part, we discuss three implementation issues for combing GSOR with the above proposed network coding scheme.

- **Reception Information Acquisition:** Network coding can only be applied when the relay node has knowledge of what its neighbors have heard, which can be obtained from two sources. One is the ACK packets received from tmp-destinations. The other is the reception reports attached to each data packet, which contain a list of packets received within a specified period.
- **Data Storage:** In our scheme, each node receives two types of packets. One is the required packets to be relayed, and the other is the overheard packets, which may be exploited to decode the encoded packet. Each receiver needs to store the required packets in its receiving buffer. In the meanwhile, each node snoops on all communications over the wireless medium and stores the overheard packets in the overhearing buffer for a limited duration. Moreover, each node has an output queue (FIFO buffer) to maintain the packets to be transferred, and periodically broadcasts reception reports to inform its neighbors which packets it has stored.
- **Data Transmission:** In the data transmission process, we utilize a sliding window to limit the number of candidate packets to be coded in the queue, where the size of the window is assumed to be static, denoted by Q. The transmission process works as follow. First, each relay node selects the unacknowledged packets from its receiving buffer which are not received by higher priority relays, and inserts such packets into the output queue. Second, the relay node codes the previous Q packets which are included in the first sending window. Third, it deletes the corresponding packets from the output queue and sends the encoded packet out. Finally, it repeats this coding process until no packets can be coded together and moves the sliding window to the next Q packets. Fig. 7 provides an example to describe this transmission procedure, where Q is set to be 7. In the sending window 1, packet 4, 5, and 7 can be coded together and delivered by one transmission. Accordingly, packet 1 and 2 can also be transmitted by one coded packet. Packet 3 and packet 6 cannot be coded, thus they are delivered by two individual transmissions. After all the packets in the sending window 1 have been transmitted successfully, the relay node will move to

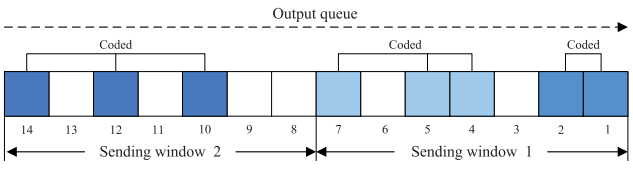


FIGURE 7. An example of output queue where gray and white squares represent packets that are codable and uncodable respectively.

sending window 2 and repeat the same transmission process.

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our proposed schemes under different network settings, e.g., PU activity, source-destination distance, transmission range, and link reliability, using Network Simulation Version 2 (NS2) extended framework [47].

A. SIMULATION SETTINGS

The network parameter settings are summarized in Table 1. We set up a CRAHN with multiple PUs and SUs, which are randomly placed in a 3000 m $\times$ 3000 m plane and assumed to be stationary throughout the simulation. We randomly choose two SUs making a source-destination node pair, associated with a constant bit rate (CBR) flow. The packet size of each flow is 512 bytes while the data rate is 10 packets per second (pps). The packet delivery ratios of the links between two neighboring nodes are based on the distance-to-delivery ratio relationship measured in [48]. In addition, the thresholds α and γ are set to be 0.05 and 2 respectively, which are the best values obtained through 1000 tests.

TABLE 1. Simulation parameters.

Number of channels	6
Number of PTs per channel	15 ($\rho_{p,i}$ is $1.7/km^2$)
PT transmission range	250 m
PT active probability	{0.7, 0.6, 0.5, 0.5, 0.3, 0.4}
Expected channel OFF time	[100 ms, 600 ms]
Number of SUs	99 (λ_s is $11/km^2$)
SU transmission range	[120 m, 320 m]
Source-destination distance	[140 m, 500 m]
SU CCC rate	1Mbps
SU data channel rate	2Mbps
Sensing time	3ms
channel switching time	100μs
retransmission time	3

In this simulation, we compare the optimal GSOR (O-GSOR) scheme, the sub-optimal (S-GSOR) scheme and the coding based S-GSOR (C-GSOR) scheme with SAOR protocol [6] and SAAR scheme [8]. Specifically, O-GSOR and S-GSOR are based on Algorithm 1 and Algorithm 2 respectively. And the C-GSOR scheme is based on Algorithm 1 and Algorithm 3. Moreover, SAOR is a well known opportunistic routing protocol coupled with spectrum sensing and sharing in multi-channel multi-hop cognitive radio networks, and SAAR is a recently published anypath routing scheme with consideration of unreliable transmission feature and uncertain spectrum availability characteristic of CR environments. We run each experiment for 50 seconds and repeat it 1000 times with different seeds to report the average value as final results.

B. THE IMPACT OF SOURCE-DESTINATION DISTANCE

In the first group of simulations, we study how the average number of transmissions varies concerning different distances between the source and the destination. We set the transmission range to be 120 m for all the CR nodes and the PUs' expected OFF time $E[T_{off}]$ to be 300 ms. We group different source-destination node pairs according to the end-to-end distances, which means the node pairs in the same group are with the same distance. We evaluate the groups whose corresponding distances are varied from 140 m to 500 m in steps of 40 m.

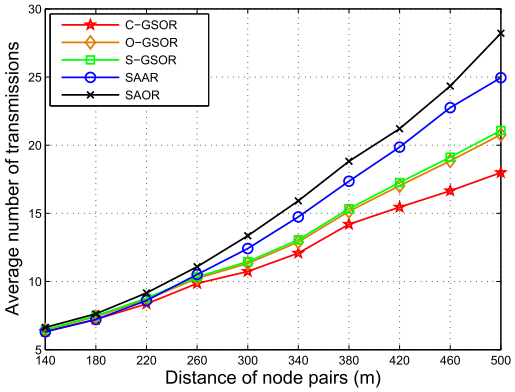


FIGURE 8. Average number of transmissions under different source-destination distances. (transmission range: 120 m, $E[T_{off}]$: 300 ms).

As shown in Fig. 8, C-GSOR always performs better than the other four schemes with less number of transmissions. With the increasing end-to-end distances, the results of the other four protocols increase more sharply than C-GSOR. The reason behind this phenomenon is that C-GSOR takes advantage of the broadcast nature of wireless networks to further reduce the number of transmissions. Specifically, when the end-to-end distance is less than 300 m, the performance gaps among C-GSOR, O-GSOR, and S-GSOR are limited. However, as the distance between the transmitter-receiver pair increases, the improvement of network coding becomes significant due to enhanced coding opportunities. Recall that our theoretical analysis shows that S-GSOR is close to the optimal solution O-GSOR. Fig. 8 confirms this statement, indicating that the gaps between O-GSOR and S-GSOR are extremely low, especially when the end-to-end distance is less than 340 m.

C. THE IMPACT OF PU ACTIVITY

We then evaluate the performance under different PU activity patterns. We set the SUs' transmission range as 120 m, and $E[T_{off}]$ varies from 100 ms to 600 ms. In this scenario, we only consider the group of node pairs whose distances are around 420 m.

As shown in Fig. 9, we observe that C-GSOR, O-GSOR, and S-GSOR provide the lower average number of transmissions than the other two methods. Furthermore, we notice

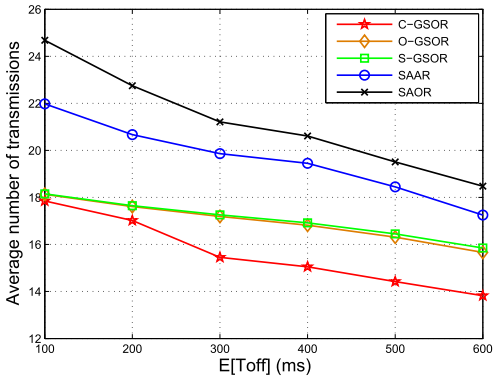


FIGURE 9. Average number of transmissions under different PU activities. (transmission range: 120 m, source-destination distance: 420 m).

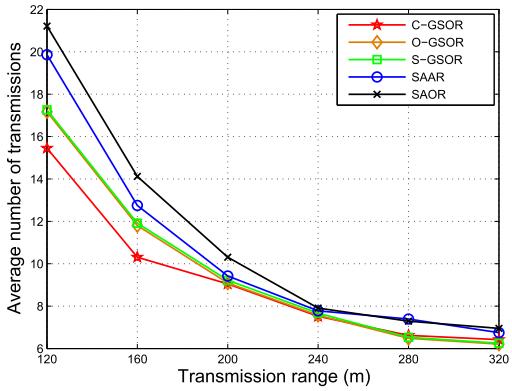


FIGURE 10. Average number of transmissions under different transmission ranges. (source-destination distance: 420 m, $E[T_{off}]$: 300 ms).

that the improvement is significant under high PU activity ($E[T_{off}]$ is below 300 ms), whereas it becomes negligible under low PU activity ($E[T_{off}]$ is above 500 ms). In CRAHNs, a small value of $E[T_{off}]$ indicates a more frequently activities of PUs, and thus the transmissions between source-destination pairs are more likely to be interrupted. Thus, the results demonstrate that our methods cope well with the dynamic changing spectrum environment, due to the step-by-step transmission policy where forwarding decisions are only based on local knowledge of the spectrum and the topology. Moreover, an interesting finding is that the coding gain becomes smaller with high PUs' activities ($E[T_{off}]$ is below 300 ms). The reason behind this result is that, with increasing loss probability induced by the PUs' frequently spectrum access, the chance of correctly decoding at the destination decreases. Moreover, Fig. 9 report that the performance of S-GSOR can lead to a sub-optimal solution, especially when the primary traffic is highly dynamic. In CRAHNs, the instability and unpredictability of channel conditions invalidate several links and nodes, even may partition the topology of the network. This case is mainly due to PU activities. Hence, the results in Fig. 9 can attest the effectiveness of our schemes when the topology is highly dynamic.

D. THE IMPACT OF TRANSMISSION RANGE

In the third group of simulations, we investigate how different transmission range values contribute to the average number of transmissions. Similar to the previous experiment, we only consider the group of node pairs whose corresponding distances are around 420 m. The PUs' expected OFF time is set to be 300 ms. We assume each SU can adjust its transmission range (using power control), which varies from 120 m to 320 m in steps of 40 m.

As can be seen from Fig. 10, the average number of transmissions drops with the increase of the transmission range for all four methods. The reason is that as the transmission range increases, the connectivity of the network is also improved. Furthermore, we notice that C-GSOR, O-GSOR, and S-GSOR can still achieve better performance than the other two protocols no matter how the transmission range varies. An interesting observation is that, if the transmission range is below 200 m, C-GSOR, O-GSOR, and S-GSOR can obtain more benefits; whereas, if the transmission range is above 200 m, the results are close to the other schemes. Moreover, the result of C-GSOR always outperforms O-GSOR and S-GSOR with less number of transmissions. Nonetheless, when the transmission range is sufficiently large, the improvement from network coding is not as substantial. The reason behind is that the opportunity of different flows intersecting at the same relay node is relatively small, which makes less coding gain. On the contrary, when the transmission range is decreased, more coding opportunities are possible to be exploited, which is attributed to produce a larger gap between C-GSOR and non-coding schemes. Fig. 10 also confirms that the performance gap between O-GSOR and S-GSOR is rather small for all cases.

E. THE IMPACT OF LINK RELIABILITY

We then proceed to evaluate the performance of various networks with different levels of reliability, which are characterized by the links' average packet delivery ratio of the network. We assign different relative weights to generate networks with varying levels of reliability, which ranges from 0.2 to 0.6 in steps of 0.1. The transmission range is fixed to be 120 m for all CR nodes, and each PU's expected OFF time is set to be 300 ms. We only consider the group of node pairs whose distances are around 300 m.

The first significant observation is that, when the average packet delivery ratio is below 0.2, O-GSOR performs slightly better than S-GSOR. The main reason for this is that, if the stopping threshold is not small enough, the criterion may cause premature termination of Algorithm 2. However, in most cases, the performance difference between O-GSOR and S-GSOR is still negligible. The second significant result from Fig. 11 is that C-GSOR, O-GSOR, and S-GSOR perform better than the other two schemes. More importantly, when the average delivery ratio is relatively low, C-GSOR, O-GSOR, and S-GSOR can achieve more superior performance. Nevertheless, when the average delivery ratio is high, less improvement is obtained. Moreover, when the average

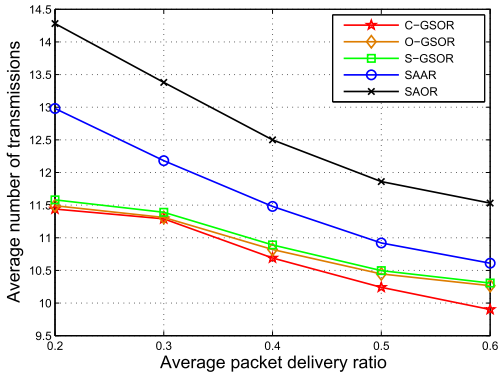


FIGURE 11. Average number of transmissions in various networks with different levels of reliability. (transmission range: 120 m, $E[T_{off}]$: 300 ms, source-destination distance: 300 m).

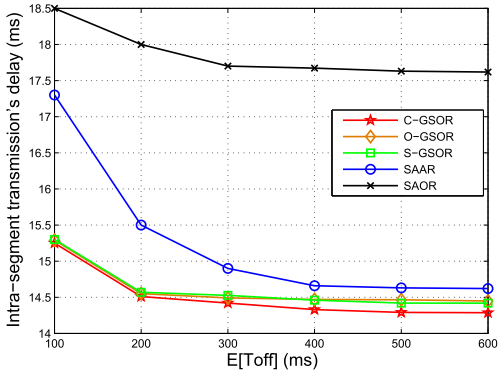


FIGURE 12. The intra-segment transmission's delay under different PU activities. (transmission range: 120 m, source-destination distance: 180 m).

delivery ratio is low, C-GSOR, O-GSOR, and S-GSOR almost achieve the same performance. The reason for this fact is that, in low-reliability networks, the desired packets may not have sufficient native packets to decode, which produces lots of retransmissions. Thus, it is not necessary to apply network coding when the average delivery ratio is extremely low. On the contrary, when the average delivery ratio is high, the substantial performance improvement is achieved. Simulation results also reflect that, in high connectivity network scenarios, the receivers would have better chance to decode the encoded packets correctly.

F. DISCUSSION ON THE DELAY PERFORMANCE

Although our main objective is to minimize the average number of transmissions, the delay performance is also an important issue in wireless networks, especially in the dynamic spectrum environment. In this subsection, we will evaluate the delay performance for two cases. One is the intra-segment transmission, where only one opportunistic route segment is required the complete the transmission between two end nodes. The other is the inter-segment transmission, where the end-to-end path comprises multiple opportunistic route segments.

1) DELAY PERFORMANCE OF THE INTRA-SEGMENT TRANSMISSION

Fig. 12. shows the delay performance of the proposed methods for the intra-segment transmission. Similar to subsection V-C, the SUs' transmission range is set to be 120 m. To compare the results under different PU activities, the value of $E[T_{off}]$ is between 100 ms and 600 ms with a step width of 100 ms. To gain insight into the intra-segment transmissions, we consider the special case in which only the node pairs whose distances are around 180 m are permitted to transmit packets in the simulated network.

As shown in the figure, C-GSOR always performs better than other schemes, especially when $E[T_{off}]$ increases above 300 ms. However, the improvement from network coding is not significant when $E[T_{off}]$ is below 300 ms. In fact, with PUs' frequent occurrences, the chance of successfully decoding the coded signals decreases and may result in the additional retransmission delay. Moreover, since the coded packets may be lost and uncorrected received, the native packets in the output queue have to wait for future transmission opportunities, which also increases the queuing delay. We also observe that O-GSOR and S-GSOR almost produce the same results for the intra-segment transmission. This finding reflects the fact that O-GSOR creates more overhead than S-GSOR, which reduces the benefits from less number of transmissions.

2) DELAY PERFORMANCE OF THE INTER-SEGMENT TRANSMISSION

Then, we evaluate the delay performance for the inter-segment transmission. In this part, we only consider the group of node pairs whose distances are around 420 m. All the other settings are the same as in the experiment for the intra-segment transmission.

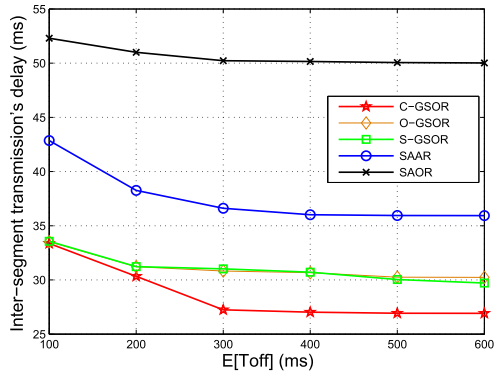


FIGURE 13. The inter-segment transmission's delay under different PU activities. (transmission range: 120 m, source-destination distance: 420 m).

Fig. 13 shows that all the five routing schemes perform better when idle state period of PU becomes longer (e.g., from 400 ms to 600 ms), as more data can be delivered during the idle state. In most cases, C-GSOR can achieve the best

performance that not only benefits from the opportunistic forwarding and step-by-step geographic routing mechanisms, but also from the network coding scheme. When the spectrum availability changes frequently (e.g., 100 ms and 200 ms), we can observe that the gap between our proposed schemes (C-GSOR, O-GSOR, and S-GSOR) and the other two methods (SAAR and SAOR) is larger than that in the relatively stable spectral usage scenarios (e.g., 500 ms and 600 ms). This result indicates that our schemes adapt better to the dynamic channel conditions. Particularly, for the case of 100 ms, a remarkable observation is that the gap between C-GSOR and O-GSOR (or S-GSOR) is negligibly small. This result confirms the fact that the coding gain becomes negligible when PUs' traffic patterns change frequently. It is intuitively obvious from this result that if the channel status changes frequently, the insufficient reception information may lead to less coding opportunities, thereby reducing the coding gain. When we compare the data of Fig. 12 and Fig. 13, we can observe that the improved performance of C-GSOR in Fig. 13 is better than that in Fig. 12. The explanation for this is two-fold. First, the inter-segment transmission involving more opportunistic route segments would reduce a larger number of transmission. Second, more flows will intersect at the relay node that creates more coding opportunities. Since O-GSOR creates more overhead than S-GSOR, when PUs' traffic patterns change frequently, S-GSOR performs better than O-GSOR. In most cases, S-GSOR and O-GSOR almost produce the same result.

G. DISCUSSION ON THE ENERGY CONSUMPTION

The current design of our schemes does not account for the power usage. However, in a real network coding based multi-path routing system, using a promiscuous mode where packet overhearing is possible, nodes may consume more power than conventional single-path routing protocols due to overhearing packets. Thus, it is necessary to measure the energy consumption of the proposed solutions. Since SAAR and SAOR do not take energy consumption into account, we compare our methods with SSR (Semi-Structure Routing) scheme [25], which is the latest published single-path routing protocol, aiming at reducing the energy consumption for CR networks. In this experiment, we assume there are 50 SUs which are equipped with limited energy supply. The transmission range is fixed to be 120 m for all the CR nodes, and each PU's expected OFF time is set to be 300 ms. We simulated a variable connection pattern, where 30 unicast flows between 30 randomly chosen source-destination pairs are created for every 10 seconds of the simulation. The duration of each simulation is 320 seconds, with a startup preparation time of 100 seconds (where no traffic is generated). Other simulation settings are the same as the previous simulations.

Fig. 14 shows how many nodes have died over time due to the lack of battery. In the first 180 seconds of simulation, with SSR scheme all the nodes can prolong their lifetime over the simulation. However, in the C-GSOR scheme, the number of nodes alive starts to decline when 120 seconds elapse. That is because, at the beginning of the simulation, performance benefits from opportunistic routing and network coding opportunities are limited. Moreover, in the first 240 seconds, C-GSOR consumes more energy than SSR because most energy is wasted in overhearing packets. More particularly, when the timeline reaches 240 seconds, an interesting observation is that C-GSOR and SSR almost perform the same. The result confirms that the additional energy overhead generated from network coding can be compensated by the energy saving from the reduced number of transmissions. The final striking observation is that O-GSOR and S-GSOR can achieve better performance than C-GSOR during the first 230 seconds. However, in the last 100 seconds, O-GSOR performs significantly worse than C-GSOR and S-GSOR. The reason for this is that involving more relay nodes would consume more energy, especially when the time horizon is long. Therefore, with limited power resources, S-GSOR is more superior than O-GSOR.

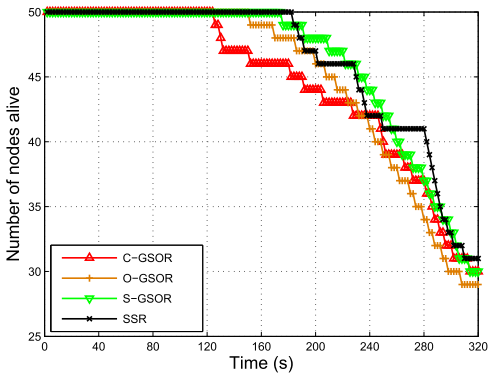


FIGURE 14. Nodes alive vs time.

VI. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we developed a distributed routing scheme GSOR to minimize the average number of transmissions per packet for CRAHNs. GSOR transmits packets through a number of intermediate destination nodes step-by-step until the ultimate destination is reached. We theoretically analyzed the properties of GSOR and proposed the optimal and sub-optimal algorithms for choosing the intermediate destination node and corresponding relay nodes. To further reduce the number of transmissions, we integrated network coding into GSOR. We constructed a coding graph to represent the relationship between the packets and the nodes, and found that our coding problem is the reduction of the maximum clique problem. We proposed a greedy algorithm to solve this coding problem and described the implementation details. Simulation results demonstrate that C-GSOR can significantly reduce the average number of transmissions and end-to-end delay, compared with existing opportunistic routing schemes in CRAHNs. The results also verify that our method can substantially improve the performance when the path from the source to the destination is relatively long, and the connectivity of the network is highly intermittent.

We will extend our work along three directions in the future. Firstly, the issue of interference cancellation and channel coefficients estimates for cross-layer design in CRAHNs is really a promising direction. In this work, we assume perfect overhearing. That is, each node guarantees that the ''unwanted'' overheard packets are correctly received. However, in real practice, interference cancellation and channel coefficients estimates are two critical practical issues on the application of network coding technique, especially in multi-channel multi-radio systems. Furthermore, it is an interesting direction to study under which conditions cooperation is the best strategy for all the involved nodes, and when it needs to be enforced. Thirdly, we may address the issue of designing an energy-efficient network coding based routing scheme for CRAHNs.

ACKNOWLEDGMENT

Part of this work was done at City University of Hong Kong. This paper was presented at the IEEE GLOBECOM Workshops (GC Wkshps), Anaheim, CA, USA, December 2012.

REFERENCES

[1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, ''NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey,'' *Comput. Netw.*, vol. 50, pp. 2127–2159, Sep. 2006.

[2] I. F. Akyildiz, W. Y. Lee, and K. R. Chowdhury, ''CRAHNs: Cognitive radio ad hoc networks,'' *Ad Hoc Netw.*, vol. 7, pp. 810–836, Jul. 2009.

[3] M. Youssef, M. Ibrahim, M. Abdelatif, L. Chen, and A. V. Vasilakos, ''Routing metrics of cognitive radio networks: A survey,'' *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 92–109, 1st Quart., 2014.

[4] S. Biswas and R. Morris, ''ExOR: Opportunistic multi-hop routing for wireless networks,'' in *Proc. ACM SIGCOMM*, 2005, pp. 133–144.

[5] B. Karp and H. T. Kung, ''GPSR: Greedy perimeter stateless routing for wireless networks,'' in *Proc. ACM MobiCom*, 2000, pp. 243–254.

[6] Y. Liu, L. X. Cai, and X. S. Shen, ''Spectrum-aware opportunistic routing in multi-hop cognitive radio networks,'' *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1958–1968, Nov. 2012.

[7] S.-C. Lin and K.-C. Chen, ''Spectrum-map-empowered opportunistic routing for cognitive radio ad hoc networks,'' *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2848–2861, Jul. 2014.

[8] J. Wang, H. Yue, L. Hai, and Y. Fang, ''Spectrum-aware anypath routing in multi-hop cognitive radio networks,'' *IEEE Trans. Mobile Comput.*, vol. 16, no. 4, pp. 1176–1187, Apr. 2017.

[9] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, ''Network information flow,'' *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[10] S. Kafaie, Y. Chen, O. A. Dobre, and M. H. Ahmed, ''Joint inter-flow network coding and opportunistic routing in multi-hop wireless mesh networks: A comprehensive survey,'' *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1014–1035, 2nd Quart., 2018.

[11] M. Ghaderi, D. Towsley, and J. Kurose, ''Reliability gain of network coding in lossy wireless networks,'' in *Proc. 27th IEEE Int. Conf. Comp. Commun. (INFOCOMM)*, Apr. 2008, pp. 2171–2179.

[12] L. Hai, H. Wang, and J. Wang, ''Instantly decodable network coding for multiple unicast retransmissions in wireless point-to-multipoint networks,'' *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6232–6243, Aug. 2016.

[13] A. Naeem, M. H. Rehmani, Y. Saleem, I. Rashid, and N. Crespi, ''Network coding in cognitive radio networks: A comprehensive survey,'' *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1945–1973, 3rd Quart., 2017.

[14] Z. Shu, J. Zhou, Y. Yang, H. Sharif, and Y. Qian, ''Network coding-aware channel allocation and routing in cognitive radio networks,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 5590–5595.

[15] X. Zhong, Y. Qin, Y. Yang, and L. Li, ''CROR: Coding-aware opportunistic routing in multi-channel cognitive radio networks,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 100–105.

[16] L. Hai, J. Wang, P. Wang, H. Wang, and T. Yang, ''High-throughput network coding aware routing in time-varying multihop networks,'' *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6299–6309, Jul. 2017.

[17] S.-C. Lin and K.-C. Chen, ''Statistical QoS control of network coded multipath routing in large cognitive machine-to-machine networks,'' *IEEE Internet Things J.*, vol. 3, no. 4, pp. 619–627, Aug. 2016.

[18] Y. Qu, C. Dong, S. Guo, S. Tang, H. Wang, and C. Tian, ''Spectrum-aware network coded multicast in mobile cognitive radio ad hoc networks,'' *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5340–5350, Jun. 2017.

[19] X. Tang and Q. Liu, ''Network coding based geographical opportunistic routing for ad hoc cognitive radio networks,'' in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2012, pp. 503–507.

[20] W. Li, X. Cheng, T. Jing, and X. Xing, ''Cooperative multi-hop relaying via network formation games in cognitive radio networks,'' in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 971–979.

[21] A. A. El-Sherif and A. Mohamed, ''Joint routing and resource allocation for delay minimization in cognitive radio based mesh networks,'' *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 186–197, Jan. 2014.

[22] M. Zareei, E. M. Mohamed, M. H. Anisi, C. V. Rosales, K. Tsukamoto, and M. K. Khan, ''On-demand hybrid routing for cognitive radio ad-hoc network,'' *IEEE Access*, vol. 4, pp. 8294–8302, 2016.

[23] A. Cammarano, F. Lo Presti, G. Maselli, L. Pescosolido, and C. Petrioli, ''Throughput-optimal cross-layer design for cognitive radio ad hoc networks,'' *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 9, pp. 2599–2609, Sep. 2015.

[24] A. El Shafie, T. Khattab, and A. S. Salem, ''Relay-assisted primary and secondary transmissions in cognitive radio networks,'' *IEEE Access*, vol. 4, pp. 6386–6400, 2016.

[25] S. Ji, M. Yan, R. Beyah, and Z. Cai, ''Semi-structure routing and analytical frameworks for cognitive radio networks,'' *IEEE Trans. Mobile Comput.*, vol. 15, no. 4, pp. 996–1008, Apr. 2016.

[26] J. Wang, H. Zhang, and X. Tang, ''Delay tolerant routing for cognitive radio vehicular ad hoc networks,'' in *Proc. IEEE 22nd Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2016, pp. 24–31.

[27] Y. Saleem, K.-L. A. Yau, H. Mohamad, N. Ramli, M. H. Rehmani, and Q. Ni, ''Clustering and reinforcement-learning-based routing for cognitive radio networks,'' *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 146–151, Aug. 2017.

[28] A. Bhorkar, M. Naghshvar, and T. Javidi, ''Opportunistic routing with congestion diversity in wireless ad hoc networks,'' *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 1167–1180, Apr. 2016.

[29] J. So and H. Byun, ''Load-balanced opportunistic routing for duty-cycled wireless sensor networks,'' *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1940–1955, Jul. 2017.

[30] M. A. Rahman, Y. Lee, and I. Koo, ''Eecor: An energy-efficient cooperative opportunistic routing protocol for underwater acoustic sensor networks,'' *IEEE Access*, vol. 5, pp. 14119–14132, 2017.

[31] M. Xiao, J. Wu, and L. Huang, ''Community-aware opportunistic routing in mobile social networks,'' *IEEE Trans. Comput.*, vol. 63, no. 7, pp. 1682–1695, Jul. 2014.

[32] X. Tang, H. Zhang, K. Zhou, and J. Wang, ''Extending access point service coverage area through opportunistic forwarding in multi-hop collaborative relay wlans,'' in *Proc. 20th IEEE Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2014, pp. 787–792.

[33] D. Zeng, S. Guo, A. Barnawi, S. Yu, and I. Stojmenovic, ''An improved stochastic modeling of opportunistic routing in vehicular CPS,'' *IEEE Trans. Comput.*, vol. 64, no. 7, pp. 1819–1829, Jul. 2015.

[34] S.-Y. R. Li, R. W. Yeung, and N. Cai, ''Linear network coding,'' *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[35] Z. Zhao, Z. Ding, M. Peng, W. Wang, and J. S. Thompson, ''On the design of cognitive-radio-inspired asymmetric network coding transmissions in MIMO systems,'' *IEEE Trans. Veh. Technol.*, vol. 64, no. 3, pp. 1014–1025, Mar. 2015.

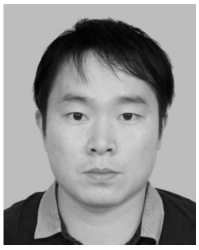
[36] W. Liang, H. V. Nguyen, S. X. Ng, and L. Hanzo, ''Adaptive-TTCM-aided near-instantaneously adaptive dynamic network coding for cooperative cognitive radio networks,'' *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1314–1325, Mar. 2016.

[37] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, ''Trading structure for randomness in wireless opportunistic routing,'' in *Proc. ACM SIGCOMM*, 2007, pp. 169–180.

[38] Y. Lin, B. Li, and B. Liang, ''CodeOR: Opportunistic routing in wireless mesh networks with segmented network coding,'' in *Proc. IEEE Int. Conf. Netw. Protocols (ICNP)*, Oct. 2008, pp. 13–22.

[39] P. Li, S. Guo, S. Yu, and A. V. Vasilakos, "CodePipe: An opportunistic feeding and routing protocol for reliable multicast with pipelined network coding," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 100–108.
[40] S. Sengupta, S. Rayanchu, and S. Banerjee, "An analysis of wireless network coding for unicast sessions: The case for coding-aware routing," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, May 2007, pp. 1028–1036.
[41] Y. Yang, Y. Liu, Q. Zhang, and L. Ni, "Cooperative boundary detection for spectrum sensing using dedicated wireless sensor networks," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.
[42] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Hoboken, NJ, USA: Wiley, 1986.
[43] D. Xue, E. Ekici, and M. C. Vuran, "Cooperative spectrum sensing in cognitive radio networks using multidimensional correlations," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1832–1843, Apr. 2014.
[44] A. W. Min, K.-H. Kim, J. P. Singh, and K. G. Shin, "Opportunistic spectrum access for mobile cognitive radios," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 2993–3001.
[45] N. Cheng, N. Zhang, N. Lu, X. Shen, J. W. Mark, and F. Liu, "Opportunistic spectrum access for CR-VANETs: A game-theoretic approach," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 237–251, Jan. 2014.
[46] B. Radunovic, C. Gkantsidis, P. Key, and P. Rodriguez, "Toward practical opportunistic routing with intra-session network coding for mesh networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 2, pp. 420–433, Apr. 2010.
[47] L. Sun, W. Zheng, N. Rawat, V. Sawant, and D. Koutsonikolas, "Performance comparison of routing protocols for cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 6, pp. 1272–1286, Jun. 2015.
[48] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker, "Complex behavior at scale: An experimental study of low-power wireless sensor networks," UCLA, Los Angeles, CA, USA, Tech. Rep. CSD-TR 02-0013, 2002.

XING TANG received the B.Eng. degree from the School of Electronic Information, Wuhan University, Wuhan, China, in 2006, and the Ph.D. degree from the Department of Computer Science, City University of Hong Kong, China, in 2015. He is currently an Assistant Professor with the School of Computer Science and Technology, Wuhan University of Technology, China. His research interests include distributed systems and wireless networks.



JUNWEI ZHOU received the Ph.D. degree from the Department of Electronic Engineering, City University of Hong Kong, China, in 2014. He is currently an Associate Professor with the School of Computer Science and Technology, Wuhan University of Technology, China. He is the Chutian Scholar of computer science and technology of Hubei province, China, in 2015. His current research interests include computer vision, biometrics, and multimedia security. He received the Outstanding Academic Performance Award and the Research Tuition Scholarships from the City University of Hong Kong.

SHENGWU XIONG received the B.Sc. degree in computational mathematics from Wuhan University, China, in 1987, and the M.Sc. and Ph.D. degrees in computer software and theory from Wuhan University, China, in 1997 and 2003, respectively. He is currently a Professor with the School of Computer Science and Technology, Wuhan University of Technology, China. His research interests include intelligent computing, machine learning, and pattern recognition.

JING WANG received the master's degree in communication and information systems from the Wuhan Research Institute of Posts and Telecommunications, China. She is currently pursuing the Ph.D. degree with the Computer School of Wuhan University, China. Her main research interests include wireless networks and mobile computing.

KUNXIAO ZHOU received the D.Phil. degree in computer science from the City University of Hong Kong in 2013. He is currently an Assistant Professor with the School of Computer Science and Network Security, Dongguan University of Technology, China. His research interests include topology control and routing in wireless networks.