Cooperative Relay Selection in Cognitive Radio Networks

Tao Jing, Shixiang Zhu, Hongjuan Li, Xiaoshuang Xing, Xiuzhen Cheng, Yan Huo, Rongfang Bie, and Taieb Znati

***Abstract*—The benefits of cognitive radio networking (CRN) have been well recognized with the emerging wireless applications in recent years. While many existing works assume that the secondary transmissions are negative interferences to the primary users (PUs), in this paper, we take secondary users (SUs) as positive potential cooperators for the PUs. In particular, we consider the problem of cooperative relay selection, in which the PUs actively select appropriate SUs as relay nodes to enhance their transmission performance. The most critical challenge for such a problem is how to select a relay efficiently. Due to the potentially large number of SUs, it is infeasible for a PU to first scan all the SUs and then pick the best. Basically, the PU transmitter intends to observe the SUs sequentially. After observing an SU, the PU needs to make a decision regarding whether to terminate its observation and use the current SU as its relay or to skip it and observe the next SU. We address this problem by using the optimal stopping theory and derive the optimal stopping rule. We also discuss the optimal observation order of the SUs and analyze the collision probability. To evaluate the performance of our proposed scheme, we compare our optimal stopping policy with the random selection policy through simulation study, and the results demonstrate the superiority of our policy. Extensive simulation study is conducted to investigate the impact of different parameters on the system performance, and the results indicate that our algorithm can satisfy different system requirements by carefully tuning the corresponding system parameters.**

***Index Terms*—Cognitive radio networks (CRNs), cooperative relay selection, optimal stopping theory, spectrum sensing order.**

I. Introduction

RECENTLY, cognitive radio networking has attracted a lot of attention from both academia and industry due to its remarkable improvement in spectrum utilization efficiency [1]–[4]. The users in a cognitive radio network (CRN) are classified into two groups: primary users (PUs) and secondary users (SUs). The PUs are authorized certain licensed spectrum bands, and the SUs can sense the unused spectrum bands, which are referred to as the *spectrum holes*, and share them with the PUs to improve spectrum utilization [5], [6].

It is observed that the direct transmissions from a primary transmitter to its primary receiver might be severely damaged due to the unstable environment in wireless communications. Thus, in this paper, we consider a cooperative relaying framework in which the PUs select the SUs that may have better channel conditions as cooperative relays to help transmit their packets. We focus on the problem of relay selection, i.e., how to efficiently find an appropriate relay that can satisfy the primary transmitter's quality of service. This problem is referred to as *cooperative relay selection* [7], [8].

The most critical concern for cooperative relay selection is *efficiency*. As the number of SUs could be large due to the rapidly growing number of mobile communication devices, it is impossible/impractical to scan/observe all the candidate relays for a primary transmitter. Thus, we propose to apply the optimal stopping theory [9], [10] for cooperative relay selection, with an objective of stopping early enough to avoid scanning all the candidate relays [11]. Apparently, different observation orders of the SU candidate relays may result in different performance when applying the stopping theory. Therefore, our second challenge is to construct an optimal observation sequence to decrease the number of candidate relays that must be scanned before stop.

Our major contributions can be summarized as follows.

- First, we formulate the problem of cooperative relay selection as an optimal stopping problem and derive the optimal stopping rule for relay selection. Our stopping criterion considers the short-term effective bit rate (instantaneous reward) and the long-term expected throughput (expected reward if all candidate relays are considered) and selects the first relay whose instantaneous reward is at least the same as the expected reward.
- Second, we investigate the impact of the observation order and obtain an optimal order that maximizes the observation efficiency. We find that a random observation order leads to an irregular and uncontrollable result in relay selection and prove that the descending order of the short-term effective data rate provides an optimal order for our stopping-theory-based cooperative relay selection mechanism.
- Third, we consider the optimal stopping policy in a multi-PU scenario and deduce the probability of packet collisions. We find out that the probability becomes smaller when the number of SUs becomes larger.

Manuscript received February 22, 2014; revised May 22, 2014; accepted July 8, 2014. Date of publication July 29, 2014; date of current version May 12, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 61172074, Grant 61272505, and Grant 61172074 and in part by the National Science Foundation under Grant CNS-1162057 and Grant CNS-1265311. The review of this paper was coordinated by Dr. X. Huang.

T. Jing, S. Zhu, X. Xing, and Y. Huo are with the School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: tjing@bjtu.edu.cn; 11120212@bjtu.edu.cn; 11111030@bjtu.edu.cn; yhuo@bjtu.edu.cn).

H. Li and X. Cheng are with the Department of Computer Science, The George Washington University, Washington, DC 20052 USA (e-mail: hongjuan@gwu.edu; cheng@gwu.edu).

R. Bie is with the College of Information Science and Technology, Beijing Normal University, Beijing 100875, China (e-mail: rfbie@bnu.edu.cn).

T. Znati is with the Department of Computer Science, University of Pittsburg, Pittsburgh, PA 15213 USA (e-mail: znati@cs.pitt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2338297

- Finally, we conduct an extensive simulation study to validate the performance of our relay selection scheme. In particular, we compare our proposed scheme with the random relay selection scheme. Moreover, we investigate the impact of different parameters and present a thorough analysis on the results.

The rest of this paper is organized as follows. The most related research is introduced in Section II. Our network model and the adopted relaying framework are illustrated in Section III. The proposed optimal stopping-policy-based cooperative relay selection scheme is detailed in Section IV. Theoretical analysis and discussions are presented in Section V. The results of performance evaluation are reported in Section VI. The conclusion is presented in Section VII.

II. Related Work

An overview of existing cooperative relaying techniques is provided in [12]. Various relay selection approaches [13]–[19] have been explored for cooperative relaying in general wireless networks. Some of them require channel-related information from all the candidate relay nodes, which is inefficient when the number of candidate relays is large. For example, partial channel state information, channel assignment, and signal-to-noise ratio (SNR) are required by the relay selection approaches proposed in [13]–[15], respectively. Moreover, in [16], the SNR thresholds for all candidate relays need to be computed. As a comparison, our approach does not require information from all candidate relay nodes as it scans the candidate SU relays one by one and stops when a suitable relay is identified.

Relay selection is sometimes jointly considered with other network functions. Guan *et al.* [17] jointly considered spectrum management and relay selection and formulated an equivalent Nash equilibrium problem. Li *et al.* [18] proposed a dynamic relay selection algorithm to achieve the tradeoff between performance and cost, where the cost refers to the energy consumption. Wang *et al.* [19] considered the location condition as a benefit in relay selection and proposed a distributed approach for relay selection and power control. Differently, our approach employs the optimal stopping theory to take into account the time to scan the candidate relays before stopping at a suitable relay with good channel quality.

The optimal stopping theory has been applied to opportunistic scheduling and spectrum sensing. Zheng *et al.* [20] proposed an optimal stopping approach for distributed opportunistic scheduling for ad hoc networks from a team game perspective. Shu and Krunz [21] tackled the sequential decision strategy to solve the channel sensing and probing problem for CRNs. Jia *et al.* [22] employed optimal stopping to consider the optimal spectrum sensing decision for a single secondary transmission pair in ad hoc CRNs. Our work is the first to apply stopping theory for cooperative relay selection in CRNs, to our best knowledge. It employs an optimal stopping rule to find out the relay with good channel quality within a short observation/scan time.

We also derive an optimal observation order to maximize the observation efficiency. Some related research has been done in [23]–[26] to investigate the optimal order for spectrum sensing in CRNs. Cheng and Zhuang [23] proposed a simple descending order based on SNR and proved that it is optimal when the user stops at the first free channel. Jiang *et al.* [24] presented a dynamic programming approach to find an optimal observation order and investigated the optimal order in some special cases. Fan and Jiang [25] proposed two suboptimal algorithms to find the optimal order in a two-user case. An approach for searching an optimal order dynamically based on reinforcement learning is proposed in [26]. Our approach employs the effective data rate to define the observation order, and we prove that this order is most efficient in terms of observation time.

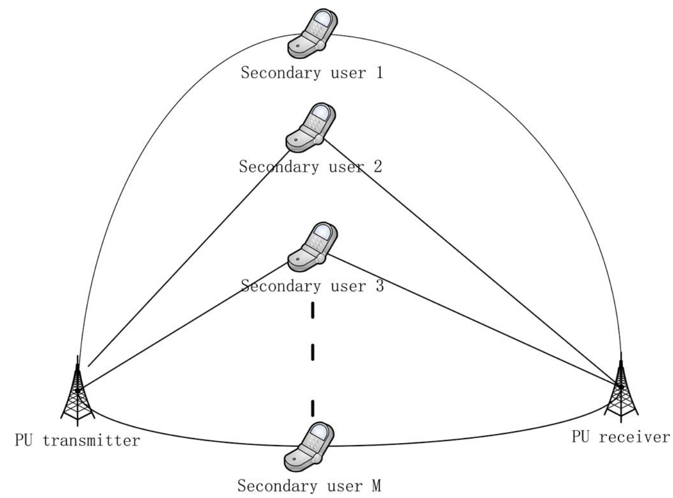

Fig. 1. Network model.

III. System Models

Here, we first illustrate the system model for our problem of cooperative relay selection in CRNs. Then, we introduce a simple cooperative relaying protocol adopted by this paper.

A. Network Model and Assumptions

There are many types of relaying techniques: decode and forward, amplify and forward (AF), quantize and forward, and compress and forward. In this paper, we adopt AF to illustrate our design. In AF, a relay node amplifies the signal of the received packets and then delivers them.

We consider a simple CRN that consists of some primary transmitters and primary receivers and a number of SUs. A typical primary transmitter, which is denoted by P_t, transmits its packets to a typical primary receiver, which is denoted by P_r, with the assistance from one of M SUs represented by s_i, $i = 1, 2, \ldots, M$, as shown in Fig. 1, where P_t and P_r form a primary transmission pair. When P_t needs to transmit packets to P_r, a free SU, which has a better channel condition compared with P_t, can be selected as a relay node by the PU pair. The M SUs, which have the ability to help transmit packets for the primary system, are called *candidate relays*, and the SU finally selected by the PU pair is called a *cooperative relay*.

It is assumed that the proposed network system is time-slotted, the cooperative relay selection is performed at each time slot, and the duration of a time slot is T. For simplicity, we

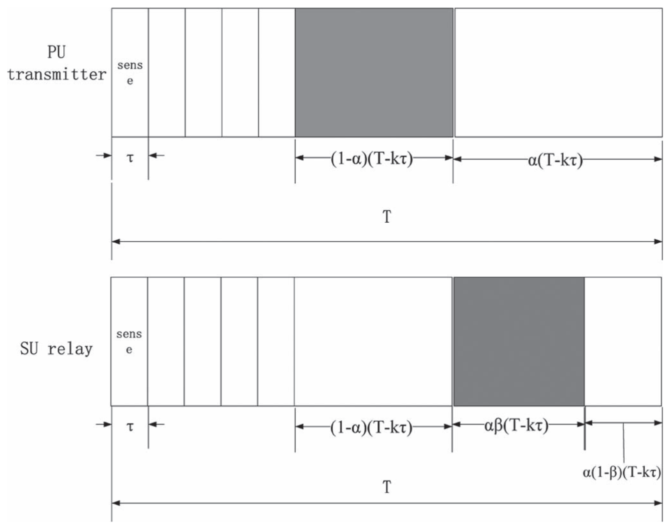

Fig. 2. Time slot structure.

further assume that each PU pair can select, at most, one SU as a cooperative relay and that each SU candidate relay node can only be selected by, at most, one PU pair. In addition, one user can only transmit or receive at one channel per time slot because each user is equipped with only a simple transceiver. For simplicity, in this paper, we first consider the network scenario where there exists only one primary pair and M SUs; then, we analyze the collision probability of multiple PU pairs simultaneously selecting the same SU as a relay.

B. Cooperative Relaying Protocol

To find a proper cooperative relay, the PU transmitter observes the candidate relay nodes to obtain their channel quality at each time slot. Then, the PU pair decides whether to select a candidate relay as the cooperative relay node based on a cooperative relay selection algorithm. It is assumed that the observation results are error free.

In this paper, we adopt the following spectrum leasing strategy that can motivate the candidate secondary relays to help the PU pair with their packet transmissions: The PUs own the spectrum resource and have the right to decide whether to lease the spectrum to candidate secondary relays in exchange for cooperation, and the candidate secondary relays have the right to decide whether to cooperate with the PU pair on the basis of the corresponding fraction of time leased for secondary transmissions.

As shown in Fig. 2, each time slot of length T is partitioned into several components. Let τ be the time needed for observing a potential relay. We assume that τ is identical for different SUs and for different time slots. Denote by $\{s_1, s_2, \ldots, s_M\}$ an observation order/sequence, which is a permutation of the SU candidate relay index set $\{1, 2, \ldots, M\}$. At the beginning of a time slot, P_t starts to observe the SU candidate relay nodes sequentially according to the observation sequence. If the reward of the kth observation satisfies a specific criterion, P_t stops at the kth SU candidate relay node and then delivers its packets (intended to the PU receiver P_r) to the secondary relay node for a fraction $(1-\alpha)$, $0 \leq \alpha \leq 1$ of time $(T - k\tau)$ shown by the shadow part in the first subgraph. For the rest of the time slot $\alpha(T - k\tau)$, the selected secondary node relays P_t's data over a β, $0 \leq \beta \leq 1$ fraction of $\alpha(T - k\tau)$ shown by the shadow part in the second subgraph and then sends its own packets during residual time $(\alpha(1-\beta)(T - k\tau))$. Note that we assume that condition $M\tau < T$ always holds. A similar cooperative relay model is adopted in [27].

IV. Optimal Stopping Policy

A. Problem Formulation

In this paper, we focus on the problem of cooperative relay selection in a CRN where a PU pair observes the SU candidate relay nodes one by one based on an observation sequence and decides whether to stop and select the SU node currently under observation as the cooperative relay node. To maximize the reward of the selection, the PU pair makes the decision based on the result of comparing the instantaneous reward and the expected reward of future observations. The instantaneous reward can be represented by the channel quality of the candidate relay being observed, and the expected reward of future observations is the reward that the PU pair can obtain if it continues observing the following candidate relays. In other words, the PU pair may stop and receive the current reward/relay or continue observing the rest of the SU relays to find a better relay. Therefore, the relay selection problem can be formulated as a sequential decision problem and can be investigated by applying the optimal stopping theory.

The stopping theory and its applications are studied in [20]. We first introduce the concept of stopping theory and then formulate the problem of cooperative relay selection as an optimal stopping problem.

Definition 1: A stopping problem is defined by two parts:

- a sequence of random variables, i.e., $X_1, X_2, \ldots$, with a known joint distribution; and
- a sequence of reward functions, i.e., $y_0, y_1(x_1), \ldots, y_\infty(x_1, x_2, \ldots)$, which are the real-valued functions of the variables previously discussed.

The objective is to find out the variable X_i in the sequence such that the reward function $y_i(x_1, x_2, \ldots, x_i)$ is maximized.

To make sure that the packets relayed by the cooperative relay node securely arrive at the destination, some conditions/restrictions should be satisfied, which can be described as follows:

$$0 < (1-\alpha)R_{\text{ps}}^r(t) \leq \alpha\beta R_{\text{sp}}^s(t) \qquad (1)$$

where $R_{\text{ps}}^r(t)$ denotes the transmission rate between the PU transmitter and the SU relay, and $R_{\text{sp}}^s(t)$ denotes the transmission rate between the SU relay and the PU receiver. The intuition behind this equation is that the amount of data transmitted from the primary transmitter should not exceed the transmitting capability of the relay. The value of α is controlled by the PU transmitter. Given α, we can derive the minimum value of β, i.e., $\beta_{\text{low}} = ((1-\alpha)R_{\text{ps}}^r(t)/\alpha R_{\text{sp}}^s(t))$. For simplicity, it is assumed that $R_{\text{ps}}^r(t) = R_{\text{sp}}^s(t)$. Thus, we have $\beta_{\text{low}} = (1-\alpha)/\alpha$, and $0.5 \leq \alpha \leq 1$ since $0 \leq \beta_{\text{low}} \leq 1$.

In a cognitive communication network, an SU may not be available to serve the PUs due to the secondary communication carried out by SUs. Therefore, the PU transmitter should examine the availability of the SU candidate relays. When the PU transmitter observes the channel condition of an SU, the SU returns a value for β. It is assumed that the SU is always in saturated transmission mode, which means that the SU candidate relay never returns a parameter $\beta = 1$. Note that the SU is available if the returned value is larger than or equal to β_{low}, and vice versa. Let θ denote the probability that the SU candidate relay is available. Then, we define Θ as the indicator function of the availability of the SU candidate relay, which is given by

$$\Theta = \begin{cases} 0, & \text{if } \beta < \beta_{\text{low}} \text{ with probability } (1-\theta) \\ 1, & \text{if } \beta \geq \beta_{\text{low}} \text{ with probability } \theta. \end{cases} \quad (2)$$

To further investigate the channel quality in our cooperative relay selection problem, we assume that the underlying channel is a flat Rayleigh fading channel, in which the instantaneous interference-plus-noise ratio (SNR) is received by the destination with an exponential distribution having a probability density function $f(\gamma) = (1/\overline{\gamma})e^{-(\gamma/\overline{\gamma})}$, where $\overline{\gamma}$ denotes the average SNR in the channel model. Then, we can model the Rayleigh fading channel as a finite-state Markov chain (FSMC), as proposed in [28]. In the FSMC, we partition the SNR into U intervals and then divide the SNR into finite-state space. Thus, the SNR thresholds are denoted by $\Upsilon = \{\gamma_1 = 0, \gamma_2, \ldots, \gamma_U, \gamma_{U+1} = \infty\}$. If an instantaneous SNR Γ is in $[\gamma_u, \gamma_{u+1})$, the channel of the SU candidate relay is said to be in state u. When the PU pair observes the channel of the candidate relay, the probability of the SU being in state u for the channel can be given by

$$q_u = \int_{\gamma_u}^{\gamma_{u+1}} f(\gamma)\, d\gamma = e^{-\frac{\gamma_u}{\overline{\gamma}}} - e^{-\frac{\gamma_{u+1}}{\overline{\gamma}}}, \qquad u = 1, \ldots, U. \quad (3)$$

Generally speaking, the achievable transmission rate is viewed as a metric for the channel quality in wireless communications. Let r_k denote the achievable transmission rate between the PU pair and the SU candidate relay node k. According to Shannon's theorem, r_k is calculated as follows:

$$r_k = W \log(1 + \gamma_k) \quad (4)$$

where W denotes the bandwidth of the spectrum in which a wireless user can transmit or receive data. Thus, the corresponding data rate, which is denoted as $R = \{r_1, r_2, \ldots, r_U\}$, can also be modeled as a discrete random variable with a distribution that is the same as the channel state

$$\Pr\{R = r_u\} = q_u, \qquad u = 1, 2, \ldots, U. \quad (5)$$

The PU pair acquires the achievable transmission rate of the channel between itself and the SU candidate relay by executing the observation procedure in relay selection. The process of observation is similar to the RTS/CTS access mechanism designed for the 802.11 technique [29]. At each observation step, the PU transmitter sends an Request-To-Send (RTS) frame to the candidate relay. Upon receiving an RTS frame, the candidate relay returns a Clear-To-Send (CTS) frame, which includes the information for calculating the achievable rate. We define $X_k = R_k\Theta$ as the valid transmission rate of the kth observation step. Then, the distribution of X_k can be calculated as follows:

$$\begin{aligned} p_0 &= \Pr\{X_k = x_0 = 0\} = (1-\theta) \\ p_u &= \Pr\{X_k = x_u = r_u\} = q_u\theta \\ &\quad \text{for} \quad 1 \leq u \leq U, \quad 1 \leq k \leq M. \end{aligned} \quad (6)$$

Now, we have known the distribution of the sequential variables. Next, we derive the reward function denoted by Y_k based on the sequential variables and the number of observation steps. First, we denote c_k as a scaling factor if the PU pair stops at the kth observed candidate relay node, which is given by

$$c_k = 1 - \frac{k\tau}{T}. \quad (7)$$

From (7), we can see that the larger the value of k, the smaller the value of c_k. In other words, the more number of SU candidate relay nodes the PU transmitter observes, the less the efficiency of the cooperative relay selection process. Therefore, the payoff after the kth observation attempt is given by

$$Y_k = \frac{X_k(T - k\tau)}{k\tau + (T - k\tau)}. \quad (8)$$

The numerator of (8) denotes the amount of data (in bits) that can be transmitted in one slot. The denominator is the total time cost for a time slot. Thus, the reward Y_k represents the average throughout the PU pair obtains at the current time slot if the pair stops after observing the kth SU candidate relay node and selects that node as the cooperative relay. By simplifying (8) previously defined, we obtain

$$Y_k = c_k X_k \quad (9)$$

which is a function of the observation variables X_k and the number of observation steps k.

After introducing the reward function, we summarize the optimal stopping problem in cooperative relay selection as follows: The PU pair receives the reward Y_k after the kth observation. Then, the PU transmitter makes a decision on whether to stop at the current candidate relay or continue to observe the next candidate relay based on the reward. Note that no recall is allowed since the channel quality is rapidly changing in CRNs due to complicated conditions such as the mobility of the users.

B. Optimal Stopping Rule

Here, we intend to solve the stopping problem previously discussed by deriving an optimal rule that decides when to stop observing the candidate relays. Before deriving the optimal rule, we first prove the existence of an optimal solution to the optimal stopping problem.

Theorem 1: An optimal stopping rule exists for the stopping problem formulated in Section IV-A.

Proof: We show that under the following two conditions, an optimal stopping rule exists (see [9, Ch. 3, Th. 1]):

$$1. \quad E\{\sup_n Y_n\} < \infty \tag{10}$$

$$2. \quad \limsup_{n\to\infty} Y_n \le Y_\infty = -\infty. \tag{11}$$

Now, we consider these two conditions separately.

From Section III-B, we can see that $M\tau < T$. This implies that the condition $c_k \subseteq (0,1)$ always holds. Moreover, the random variable X_n is independent and identically distributed and takes a finite number of values. Therefore, it holds that $E\{\sup_n Y_n\} < \infty$.

It is easy to see that $c_n = (1 - (n\tau/T)) \to \infty$ if $n \to \infty$. Since the random variable X_n takes a finite number of values, it holds that $\lim\sup_{n\to\infty} Y_n = \sup_{n\to\infty} c_n X_n \to \infty$. ■

Next, we derive the *optimal stopping rule* as the solution to the stopping problem. We formulate our solution approach as a backward induction. Denote by $V_j^{(M)}(x_1, x_2, \ldots, x_j)$ the maximum return that the PU transmitter can obtain after observing the jth candidate relay, which is given by

$$\begin{aligned} &V_j^{(M)}(x_1, x_2, \ldots, x_j) \\ &\quad = \max\Big\{ y_j(x_1, x_2, \ldots, x_j), \\ &\qquad\qquad E\Big\{ V_{(j+1)}^{(M)}\left(x_1, x_2, \ldots, x_j, X_{(j+1)}\right) \\ &\qquad\qquad\qquad \times |X_1 = x_1, X_2 = x_2, \ldots, X_j = x_j \Big\}\Big\} \end{aligned} \tag{12}$$

where $y_j(x_1, x_2, \ldots, x_j)$ represents the instantaneous reward after the kth observation, and $E\{V_{(j+1)}^{(M)} | X_1 = x_1, X_2 = x_2, \ldots, X_j = x_j\}$ represents the expected reward achieved by proceeding to observe the next candidate relay.

It is optimal to stop at step j if $V_j^{(M)}(x_1, x_2, \ldots, x_j) = y_j(x_1, x_2, \ldots, x_j)$, and to continue observing the candidate relays otherwise. In other words, the optimal stopping rule is achieved if the following condition holds:

$$\begin{aligned} y_j(x_1, x_2, \ldots, x_j) > E\Big\{ &V_{(j+1)}^{(M)}\left(x_1, x_2, \ldots, x_j, X_{(j+1)}\right) \\ &\times |X_1 = x_1, X_2 = x_2, \ldots, X_j = x_j \Big\}. \end{aligned} \tag{13}$$

To better understand the backward induction, we define Z_{M-j} to be the expected reward $E\{V_{(j+1)}^{(M)}\}$ if the PU pair proceeds to observe the next SU relay. That is

$$\begin{aligned} Z_{M-j} = E\Big\{ &V_{(j+1)}^{(M)}\left(x_1, x_2, \ldots, x_j, X_{(j+1)}\right) \\ &\times |X_1 = x_1, X_2 = x_2, \ldots, X_j = x_j \Big\}. \end{aligned} \tag{14}$$

The channels of different SU candidate relays are mutually independent since the SUs behave independently. Thus, the set $\{X_1, X_2, \ldots, X_M\}$ for the SU candidate relays are also mutually independent, indicating that $V_j^{(M)}$ only depends on X_j and Z_{M-j}. Thus, we can conclude that Z_{M-j} is a constant that only depends on $M - j$, the remaining number of steps to continue. Note that Z_j can be computed as follows:

$$Z_o = -\infty \tag{15}$$

$$Z_1 = E[Y_M] = E[c_M X_M] \tag{16}$$

$$= c_M \sum_{k=0}^{U} x_k p_k \tag{17}$$

and for $j \ge 1$

$$\begin{aligned} Z_{j+1} &= E \max\{Y_{M-j}, Z_j\} \\ &= \sum_m c_{M-j} x_m p_m + \sum_n Z_j p_n \end{aligned} \tag{18}$$

where $m \in \{k \,|\, c_{M-j} x_k > Z_j, k = 0, 1, \ldots, U\}$, $n \in \{k | c_{M-j} x_k < Z_j, k = 0, 1, \ldots, U\}$, subject to

$$\begin{cases} 0 \le m \le U \\ 0 \le n \le U \\ m + n = U. \end{cases} \tag{19}$$

The description of the optimal stopping rule is presented in Algorithm IV-B. The PU pair observes an SU candidate relay node according to the observation sequence S and obtains an instantaneous reward y_k after the kth observation (line 4). Then, the PU pair compares the value of y_k with the value of Z_{M-k} and decides to stop at step k if $y_k > Z_{M-k}$ and to continue observing the next SU candidate relay node otherwise (lines 5–9). Note that if the PU pair observes the last SU candidate relay in the observation sequence (that means the first $M - 1$ SU candidate relays do not satisfy the quality requirement of the PU pair), it has to select the Mth SU node as the cooperative relay, regardless of the value of the instantaneous reward y_M (line 12), which is the worst case of relay selection.

Algorithm IV-B: The Optimal Stopping Rule

1: Construct the observation sequence $S = \{s_1, s_2, \ldots, s_M\}$;
2: Start observation for a cooperative relay from s_1;
3: **for** $k \leftarrow 1$ to $M - 1$ **do**
4:     Compute the achievable transmission rate r_k after the kth observation and the reward y_k given by (9);
5:     Compare the value of y_k and the expected reward Z_{M-k} given by (14);
6:     **if** $y_k < Z_{M-k}$ **then**
7:         Proceed to observe the next SU candidate relay;
8:     **else**
9:         Stop at the current step and select the kth SU node as the cooperative relay;
10:     **end if**
11: **end for**
12: Select the Mth SU node s_M as the cooperative relay;

V. Performance Analysis

A. Relay Observation Order for the Optimal Stopping Policy

From Algorithm IV-B, we can easily figure out that it is not only the observation reward but the observation sequence that also affects the optimal rule. In other words, by constructing a more effective observation order before executing the relay selection procedure, the PU pair can find a high-quality SU relay node more efficiently. Motivated by this fact, we investigate an optimal observation order strategy for the optimal stopping policy in this section.

Some work has been done to search for an optimal order of the optimal stopping problem. In [30] finding an optimal order to maximize the expected reward $V\{X_1, X_2, \ldots, X_M\} = \sup EX_\tau$ is the focus of the work. In a different way, we mainly investigate the optimal observation order from the perspective of the observation efficiency, because the PU pair intends to finish the relay selection process as soon as possible such that it can have more time to transmit its data. The concept of *efficiency* is defined as follows.

Definition 2: If the PU pair stops and selects a suitable cooperative relay more quickly when observing the SU candidate relays based on an observation order denoted by S_1 compared with another observation order denoted by S_2, we say that the order S_1 is more **efficient** than the order S_2.

First, we consider a random observation order strategy in which the PU pair constructs the observation sequence randomly at every time slot. Since the observation variable set $\{X_1, X_2, \ldots, X_M\}$ is independent for each time slot, the efficiency of this order strategy is uncontrollable. Due to the poor performance of the random observation order strategy, we take into consideration an *Intuitive Order*, the concept of which is similar to that in [24]. It is defined as follows.

Definition 3: If an observation order $S = \{s_1, s_2, \ldots, s_M\}$ is constructed according to the descending order of X_k, the observation order is called an **Intuitive Order**.

Then, we prove the optimality of the Intuitive Order in this paper.

Theorem 2: The Intuitive Order is an optimal order for the optimal stopping rule proposed in Section IV-B.

Proof: We prove this theorem by contradiction. Suppose that there exists an optimal order $S^* = \{s_1^*, s_2^*, \ldots, s_k^*, s_{k+1}^*, \ldots, s_M^*\}$ and an integer k, where $1 \le k \le M-1$, such that $x_{k+1}^* > x_k^*$ in S^*. Then, the instantaneous reward after the kth observation according to this optimal order is y_k^*. We can get a new observation order by switching the positions of the SU candidate relays s_k^* and s_{k+1}^*. Let the new observation order be $S^{\text{new}} = \{s_1^*, s_2^*, \ldots, s_{k+1}^*, s_k^*, \ldots, s_M^*\}$. Then, the instantaneous reward after the kth observation according to the new order is y_k^{new}. We have

$$y_k^{\text{new}} = c_k x_{k+1}^* > y_k^* = c_k x_k^*. \tag{20}$$

Since the expected reward achieved when the PU pair proceeds to observe after the kth observation, i.e., Z_{M-k}, only depends on the remaining number of steps to continue, the condition $Z_{M-j}^* = Z_{M-j}^{\text{new}}$ holds. Thus, the PU pair observes that the SU candidate relay based on the new order has a larger probability of stopping after the kth observation than the optimal order. In other words, the new order is more efficient than the optimal order S^*, which contradicts the optimality of the optimal order S^*. ■

B. Probability of Packet Collision

Now, we consider the multi-PU pair scenario where each PU pair observes and selects the SU relay node based on its own relay observation order independently. Therefore, we may face the problem that a particular SU relay node is observed available and selected by two or more PU pairs simultaneously. In this case, these PU pairs would transmit their packets simultaneously to the same SU cooperative relay, which may cause packet collisions. Here, we intend to investigate the probability of packet collisions. First, we derive the probability of packet collisions in a scenario of two PU pairs in which the two PU pairs execute the relay selection algorithm independently.

Let Φ_k denote the probability that the PU pair decides to stop and selects the current SU candidate relay after the kth observation, which can be computed as follows:

$$\begin{aligned}\Phi_k &= \Pr\{Y_k \ge Z_{m-k}\} = \Pr\{c_k X_k \ge Z_{m-k}\} \\ &= \Pr\left\{X_k \ge \frac{Z_{m-k}}{c_k}\right\} = \sum_i p_i\end{aligned} \tag{21}$$

where $i \in \{j | x_j \ge (Z_{m-k}/c_k), 0 \le j \le U\}$ for $1 \le k \le M-1$, and ϕ_0 is defined as $\phi_0 = 0$. Let $S^n = \{s_1^n, s_2^n, \ldots, s_M^n]\}$ denote the observation orders of the nth PU pair, where $n \in \{1, 2\}$. It is assumed that the two PU pairs construct their observation orders independently and that each pair has no knowledge about the order of the other. Denote by $P_{c(2)}$ the probability of packet collisions in this two PU scenario, which is given by

$$P_{c(2)} = P_{c(2)}^1 + P_{c(2)}^2 + \cdots + P_{c(2)}^M \tag{22}$$

where $P_{c(2)}^k = \Pr\{$a packet collision occurs at the observation step $k\} = \Pr\{s_1^1$ is observed unsatisfied, s_1^2 is observed unsatisfied, s_2^1 is observed unsatisfied, s_2^2 is observed unsatisfied, ..., $s_k^1 = s_k^2$, and both are observed satisfied$\}$ and can be computed as follows:

$$\begin{aligned}P_{c(2)}^k &= \frac{M(M-1)!(M-1)!}{M!M!} \prod_{i=0}^{k-1} (1-\phi_i)^2 \phi_k^2 \\ &= \frac{1}{M} \prod_{i=0}^{k-1} (1-\phi_i)^2 \phi_k^2\end{aligned} \tag{23}$$

where $1 \le k \le M-1$. Note that the PU pair must stop and select the Mth SU candidate relay if its observation reaches the last SU candidate relay. Then, $P_{c(2)}^M$ is given by

$$P_{c(2)}^M = \frac{1}{M} \prod_{i=0}^{M-1} (1-\phi_i)^2. \tag{24}$$

Thus, the value of ϕ_M can be defined as $\phi_M = 1$, and the probability of packet collision $P_{c(2)}$ can be calculated by

$$P_{c(2)} = \sum_{k=1}^{M} \frac{1}{M} \prod_{i=0}^{k-1} (1-\phi_i)^2 \phi_k^2. \tag{25}$$

Next, we consider the probability of packet collisions in a multi-PU scenario in which N PU pairs observe and select the SU candidate relay nodes according to their observation orders independently, where $N > 2$. Note that the probability that three PU pairs stop on the same SU relay node at the kth step denoted by $P_{c(3)}^k$ can be similarly computed as (23) as follows:

$$\begin{aligned} P_{c(3)}^k &= \frac{M(M-1)!(M-1)!(M-1)!}{M!M!M!} \cdot \prod_{i=0}^{k-1} (1-\phi_i)^3 \phi_k^3 \\ &= \frac{1}{M^2} \prod_{i=0}^{k-1} (1-\phi_i)^3 \phi_k^3. \end{aligned} \tag{26}$$

Thus, the probability of packet collision for three PU pairs, which is denoted by $P_{c(3)}$, is given by

$$P_{c(3)} = \sum_{k=1}^{M} \frac{1}{M^2} \prod_{i=0}^{k-1} (1-\phi_i)^3 \phi_k^3. \tag{27}$$

Similarly, the probability of packet collision for n PU pairs denoted by $P_{c(k)}$ is given by

$$P_{c(n)} = \sum_{k=1}^{M} \frac{1}{M^{(n-1)}} \prod_{i=0}^{k-1} (1-\phi_i)^n \phi_k^n. \tag{28}$$

Next, we consider the probability of packet collisions in a multi-PU scenario where N PU pairs observe and select the SU candidate relay nodes according to their observation orders independently, where $N > 2$. Note that the probability of packet collisions at the N PU scenario, which is denoted by P_c^N, is given by

$$P_c^N = \sum_{k=2}^{N} \binom{N}{k} P_{c(k)}. \tag{29}$$

Thus, the probability that there is no collision in an N-pair scenario is $1 - P_c^N$. It is easy to see that the more PU pairs in the scenario, the larger the probability of packets collisions. This conclusion is in accord with the practical situation.

VI. Performance Evaluation

Here, we evaluate the performance of our proposed optimal stopping policy by an extensive simulation study. It is assumed that the duration of a time slot in our system is 0.2 ms. Moreover, the achievable transmission rate R_k of the kth SU candidate relay does not change within one slot. We divide the finite-state space of SNR received by the receiver into $U = 20$ intervals. The average SNR $\overline{\gamma}$ in the Rayleigh fading channel model is set to be 30 dB. The bandwidth W is set to be 1 MHz.

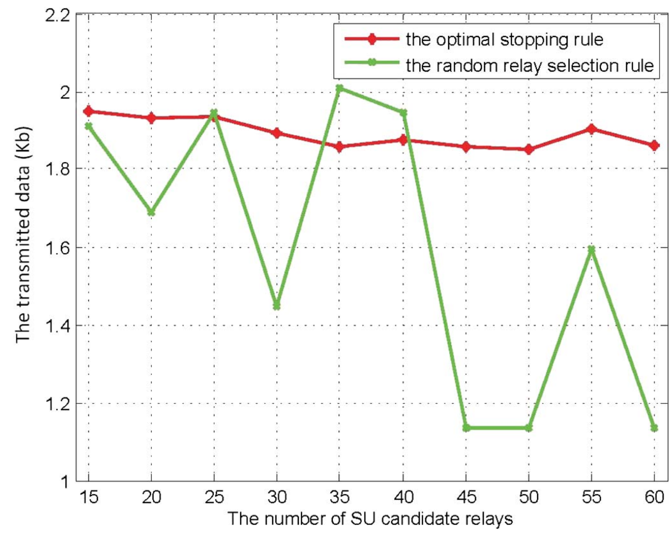

Fig. 3. Transmitted data in bits versus the number of candidate relays.

The numerical results reported in this section are averaged over 100 independent runs.

First, we describe and analyze a simple random relay selection scheme for performance comparison between the optimal stopping policy and the nonoptimal stopping policy in this section. To provide a deep insight into the optimal stopping problem in SU relay selection, we study the impact of the time duration needed for each observation (termed observation duration) τ and the parameter α on the system performance in terms of the number of observation steps and the average reward for the PU pair. We also assess the collision probability influenced by different parameters.

A. *Performance Comparison Between Our Optimal Scheme and the Random Scheme*

Here, we compare the performance of the proposed scheme with that of random relay selection in terms of the average amount of transmitted bits in one time slot. The observation time τ is set to be 3 μs, and the number of SU candidate relays ranges from 15 to 60.

In the random relay selection scheme, the PU pair randomly selects an SU candidate relay regardless of the achievable transmission rate. Since there is no time for relay selection, the whole time slot T is used for transmitting packets. Differently, the transmission time in our proposed scheme is $T - k\tau$ if the PU pair stops at the kth observation step. Fig. 3 shows the average amount of transmitted data for the two policies under different network scales. We can see that although the transmission time of the random relay selection scheme is larger than that of our proposed scheme, its average amount of transmitted data is much less for most network sizes. Moreover, the amount of transmitted data in random relay selection changes sharply and irregularly with the increase in the network size. On the contrary, our scheme demonstrates a stable transmission status regardless of the number of SU candidate relays. Thus, our optimal stopping policy not only outperforms the random relay selection policy in terms of the overall transmitted data but maintains a stable status as well.

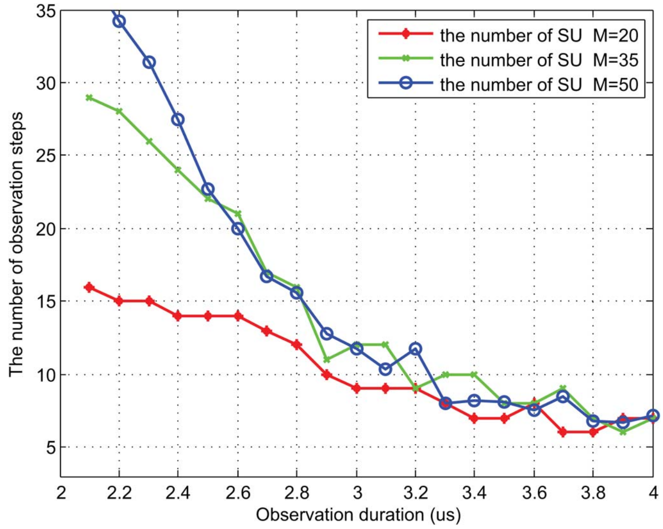

Fig. 4. Number of observation steps versus the observation duration.

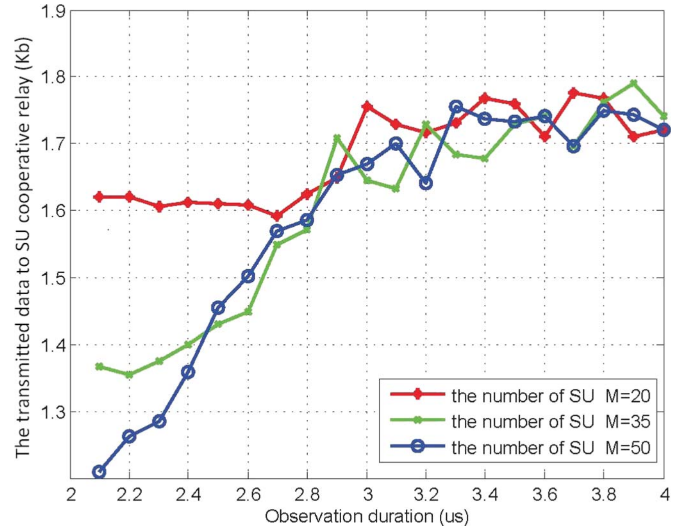

Fig. 6. Transmitted data versus the observation duration.

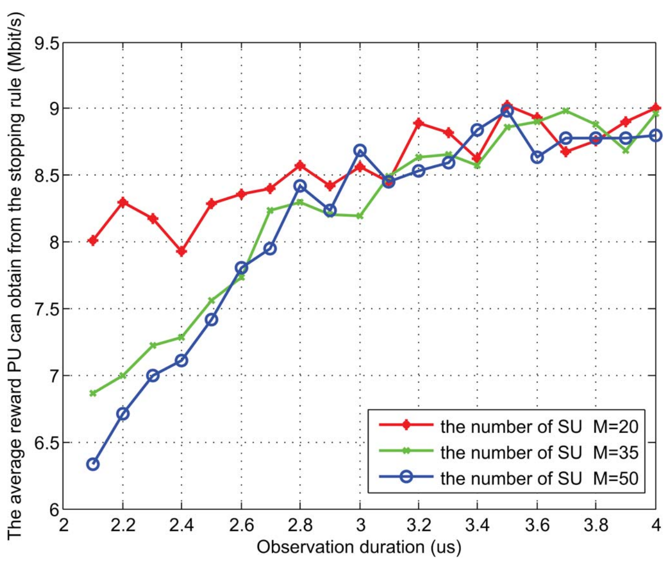

Fig. 5. Average reward for the PU pair versus the observation duration.

B. Impact of Observation Duration τ

Here, we set the fraction α to be 0.8. In Figs. 4–6, we consider different metrics under different network scales. Apparently, the observation time τ influences the relay selection performance. Fig. 4 shows that the number of observation steps decreases with the increase of the time duration needed for each observation. In other words, the larger the τ, the smaller the number of observation steps. This is because the value of τ represents the cost of observing one SU candidate relay, the PU needs to stop the observation as soon as possible to avoid generating a large cost. On the contrary, when the value of τ is small, the cost for observation is low, and the PU tends to observe more SUs to find a better cooperative relay. We also notice that the number of observation steps is larger with a larger network size, since the PU pair is provided with more choices when the number of candidate relays is larger. One important common feature among those in Figs. 4–6 is that the three curves for three different network sizes start to converge when the observation duration becomes large enough, for example, $\tau > 3$ μs.

In Fig. 5, we can see that the average reward obtained by a PU pair increases with the time duration for each observation. Jointly considering the relationship between the observation duration and the number of observation steps, we notice that if there are fewer observation steps, the larger the average reward a PU pair can obtain. This fact motivates the PU pair to construct a more efficient observation order and find the optimal cooperative relay as soon as possible. Furthermore, the fact that the number of observation steps in Fig. 4 is no more than the number of SU candidate relays in the whole network turns out to be another proof of the correctness of Theorem 1. Note that when the observation duration increases up to 3 μs or more, the number of observation steps and the average reward tend to reach a steady range. This implies that the proposed optimal stopping policy results in a steady state of the relay selection system.

In Fig. 6, we can see that the amount of data transmitted to the cooperative relays increases with the increase in the time duration for each observation τ and tends to reach a constant value. Note that although a larger observation duration implies a smaller number of observation steps as shown in Fig. 4, the time cost for each observation is larger. Thus, it is crucial to choose a right value of τ to address this tradeoff. In this simulation study, a proper value of 3 μs is selected.

We are also interested in the problem regarding how the observation duration influences the probability of packet collisions in a multi-primary pair scenario. Since the probability that a specific candidate relay is observed by two PU pairs simultaneously is identical at each step and the distribution of the variable X_k is available, the probability of the collision occurs in a specific step is a constant, and $P_{c(2)}$ can be viewed as a steady probability. This can be verified in Fig. 7. The probability of packet collision changes a little regardless of the observation duration. We also notice that different candidate relays generate different collision probabilities. A deeper analysis of this phenomenon is presented in the following section.

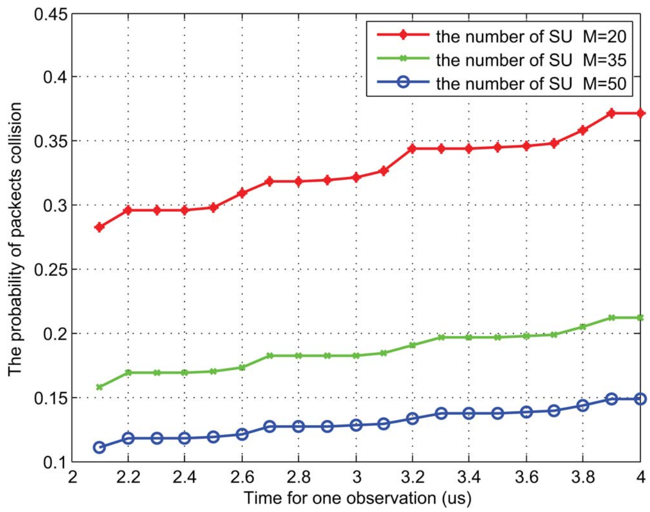

Fig. 7. Collision probability versus the observation duration.

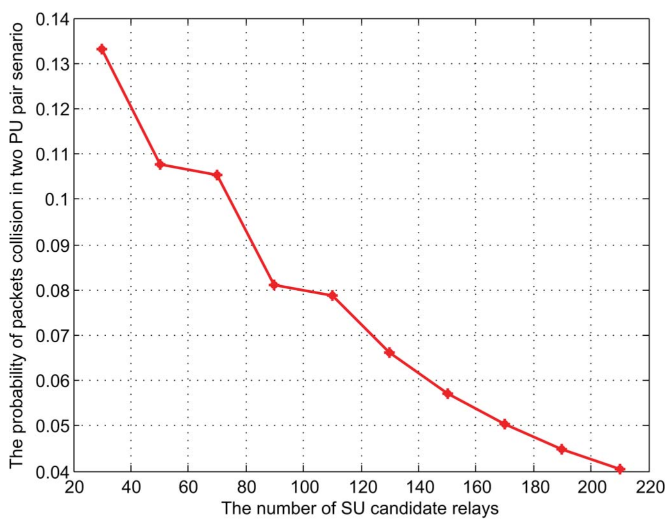

Fig. 8. Collision probability versus the number of candidate relays.

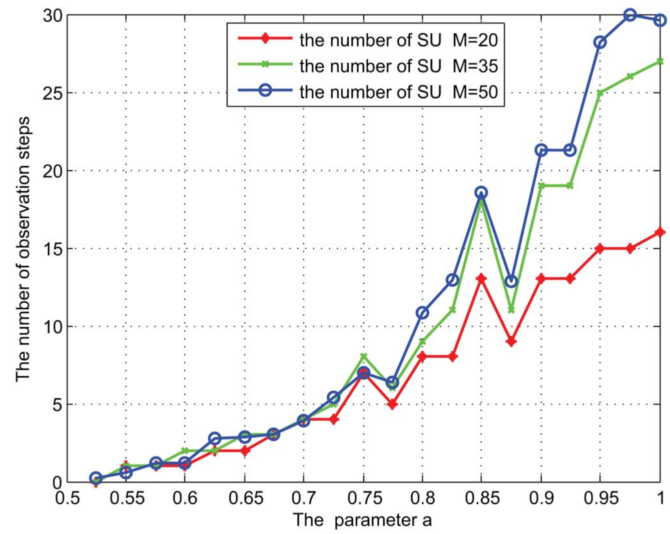

Fig. 9. Number of observation steps versus parameter α.

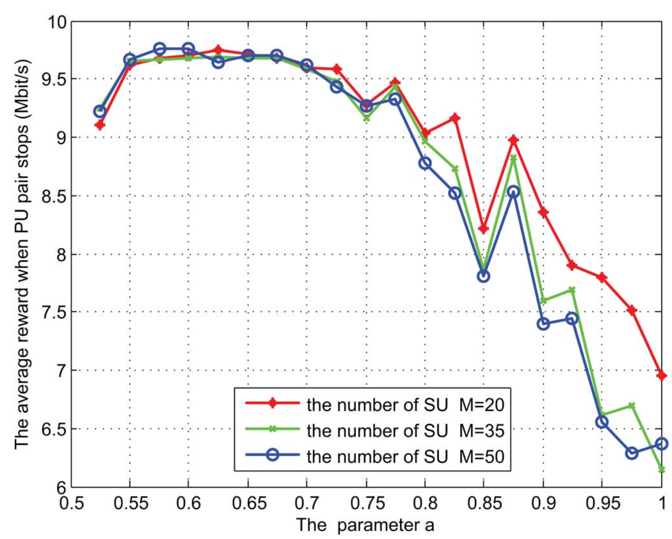

Fig. 10. Average reward versus parameter α.

C. Impact of the Number of Candidate Relays

In Figs. 4 and 5, one can see that the number of SU candidate relays has an impact on the performance of relay selection. Next, we study the impact of the number of SU candidate relays on the collision probability. The results in Fig. 8 reveal that with the growth of the number of SU candidate relays, the collision probability distinctly decreases. This descending trend can be easily understood. When the PU has more SU candidate relays to choose from, it has a smaller probability to select a candidate relay that is also simultaneously chosen by another PU. As a function of m, the collision probability decreases more apparently during the interval (0, 80).

D. Impact of the Parameter α

To further investigate the availability of the SU candidate relays, we evaluate the system performance under the situation where parameter α regularly changes. The relationship between parameter α and the availability of the SU candidate relays can be obtained in Section III-B. That is, the larger the parameter α, the smaller the threshold of parameter β, leading to a higher availability of the SU candidate relays. This relationship can be observed in Fig. 9, which demonstrates that a larger α, which implies a higher availability of an SU candidate relay, results in more number of SU candidate relays that can help the PU pair transmit packets. Therefore, we conclude that a PU pair is willing to spend more time to find a better SU candidate relay.

Similar results can be obtained in Fig. 10. When the number of observation steps grows, the scaling factor c_k quickly decreases. This explains why the average reward decreases with the growth of the observation steps. When the parameter $\alpha \leq 0.8$, the number of SU relays has little impact on the system performance.

VII. Conclusion

In this paper, we have proposed an optimal stopping policy to solve the problem of cooperative relay selection in CRNs. In our system, a PU pair observes the SU candidate relays in certain order and selects one as their cooperative relay if the transmission requirement of the SU is satisfied. We formulated an optimal stopping problem and proved the existence of the optimal solution to the stopping problem. Then, we derived an optimal stopping policy to find the optimal solution. We also defined an intuitive observation order and proved its optimality from the aspect of efficiency. The superiority of our scheme was demonstrated by comparing with the random relay selection scheme through numerical simulation. Our simulation results also reveal the impact of different parameters on the system performance.

In our future research, we plan to jointly consider relay selection and channel assignment to enhance the dynamic spectrum access efficiency. Moreover, we intend to extend our current investigation to the problem of selecting multiple relays (single-hop or multiple-hop) for each primary transmitter.

REFERENCES

[1] 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*, 2013, pp. 971–979.

[2] W. Li *et al.*, "Spectrum assignment and sharing for delay minimization in multi-hop multi-flow CRNS," *IEEE J. Sel. Areas Commun., Special Issue on Cognitive Radio*, vol. 31, no. 11, pp. 2483–2493, Mar. 2013.

[3] X. Xing, T. Jing, Y. Huo, H. Li, and X. Cheng, "Channel quality prediction based on Bayesian inference in cognitive radio networks," in *Proc. IEEE INFOCOM*, 2013, pp. 1465–1473.

[4] S. Yoon *et al.*, "Quicksense: Fast and energy-efficient channel sensing for dynamic spectrum access networks," in *Proc. IEEE INFOCOM*, 2013, pp. 2247–2255.

[5] M. Song, C. Xin, Y. Zhao, and X. Cheng, "Dynamic spectrum access: From cognitive radio to network radio," *IEEE Wireless Commun.*, vol. 19, no. 1, pp. 23–29, Feb. 2012.

[6] T. Jing, X. Chen, Y. Huo, and X. Cheng, "Achievable transmission capacity of cognitive mesh networks with different media access control," in *Proc. IEEE INFOCOM*, Mar. 25–30, 2012, pp. 1764–1772.

[7] X. Cheng, D.-Z. Du, L. Wang, and B. Xu, "Relay sensor placement in wireless sensor networks," *Wireless Netw.*, vol. 14, no. 3, pp. 347–355, Jun. 2008.

[8] X. Cheng, X. Huang, D. Li, W. Wu, and D.-Z. Du, "A polynomial-time approximation scheme for the minimum-connected dominating set in *ad hoc* wireless networks," *Networks*, vol. 42, no. 4, pp. 202–208, Mar. 2003.

[9] T. Ferguson, Optimal Stopping and Applications. [Online]. Available: http://doi.acm.org/10.1145/1614320.1614325

[10] Y. Liu and M. Liu, "To stay or to switch: Multiuser dynamic channel access," in *Proc. IEEE INFOCOM*, 2013, pp. 1249–1257.

[11] T. Jing, S. Zhu, H. Li, X. Cheng, and Y. Huo, "Cooperative relay selection in cognitive radio networks," in *Proc. IEEE INFOCOM-Mini*, 2013, pp. 175–179.

[12] W. Elmenreich *et al.*, "Building blocks of cooperative relaying in wireless systems," *Elektrotechnik und Informationstechnik*, vol. 125, no. 10, pp. 353–359, Oct. 2008.

[13] S. I. Hussain, M.-S. Alouini, M. Hasna, and K. Qaraqe, "Partial relay selection in underlay cognitive networks with fixed gain relays," in *Proc. IEEE VTC-Spring*, May 2012, pp. 1–5.

[14] J. Jia, J. Zhang, and Q. Zhang, "Cooperative relay for cognitive radio networks," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 2304–2312.

[15] C.-T. Chou, J. Yang, and D. Wang, "Cooperative MAC protocol with automatic relay selection in distributed wireless networks," in *Proc. 5th Annu. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, Mar. 2007, pp. 526–531.

[16] J. Ai, A. Abouzeid, and Z. Ye, "Cross-layer optimal policies for spatial diversity relaying in mobile *ad hoc* networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 2930–2939, Aug. 2008.

[17] Z. Guan, T. Melodia, D. Yuan, and D. A. Pados, "Distributed spectrum management and relay selection in interference-limited cooperative wireless networks," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, 2011, pp. 229–240.

[18] Y. Li, P. Wang, D. Niyato, and W. Zhuang, "A dynamic relay selection scheme for mobile users in wireless relay networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 256–260.

[19] B. Wang, Z. Han, and K. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using buyer/seller game," in *Proc. IEEE INFOCOM*, May 2007, pp. 544–552.

[20] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad-hoc communications: An optimal stopping approach," in *Proc. 8th ACM Int. Symp. Mobile ad hoc Netw. Comput.*, 2007, pp. 1–10.

[21] T. Shu and M. Krunz, "Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors," in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, 2009, pp. 37–48.

[22] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 106–117, Jan. 2008.

[23] H. T. Cheng and W. Zhuang, "Simple channel sensing order in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 676–688, Apr. 2011.

[24] H. Jiang, L. Lai, R. Fan, and H. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[25] R. Fan and H. Jiang, "Channel sensing-order setting in cognitive radio networks: A two-user case," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4997–5008, Nov. 2009.

[26] A. Mendes, C. Augusto, M. da Silva, R. Guedes, and J. de Rezende, "Channel sensing order for cognitive radio networks using reinforcement learning," in *Proc. IEEE 36th Conf. LCN*, Oct. 2011, pp. 546–553.

[27] K. Khalil, M. Karaca, O. Ercetin, and E. Ekici, "Optimal scheduling in cooperate-to-join cognitive radio networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 3002–3010.

[28] H. S. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.

[29] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[30] D. Gilat, "On the best order of observation in optimal stopping problems," *J. Appl. Probab.*, vol. 24, no. 3, pp. 773–778, Sep. 1987. [Online]. Available: http://www.jstor.org/stable/3214107

Authors' photographs and biographies not available at the time of publication.