Best Relay Selection for Underlay Cognitive Radio Systems with Collision Probability Minimization

Yahya H. Ezzeldin†, Ahmed Sultan†, Moustafa Youssef*
†Department of Electrical Engineering, Alexandria University, Alexandria, Egypt
*Wireless Research Center, Egypt-Japan Univ. of Sc. & Tech. (E-JUST) and Alex. Univ., Alexandria, Egypt

Abstract—We consider an underlay relay-assisted two-hop cognitive link which employs best relay selection to enhance the performance of secondary transmissions while keeping the interference at the primary receiver below a specified level. Best relay selection is implemented using countdown timers without cooperation between the relays. Assuming channel knowledge at the relays, the best relay is determined by the maximum signal-to-noise ratio (SNR) at the secondary receiver. Due to the timer-based operation, collisions may occur and the timer design is optimized to minimize the collision probability. The performances of the amplify-and-forward and decode-and-forward modes of operation are compared. Closed form expressions for the secondary outage probability are derived for both schemes. Numerical results provide insight into the system operation and how the outage probability depends on the various system parameters.

I. INTRODUCTION

Cognitive Radios (CR) is a promising technology for solving the scarcity in bandwidth due to the inefficient policies of fixed spectrum assignments [1]. As pointed out in [2] and other recent work, the solution for this inefficiency in bandwidth utilization is to use dynamic and adaptive spectrum access strategies. Nodes need to be able to tap into underutilized licensed bands when there is sparse or no activity and leave the band when a primary user begins to access this band.

A core paradigm in cognitive radio transmission is underlay transmission. Underlay systems are able to co-exist with primary networks given that they abide by pre-set transmission constraints. Specifically, the secondary users (SUs) of a cognitive network are allowed to transmit concurrently with the primary users so long as the interference they cause at the primary receive nodes do not exceed a certain threshold [3].

Cooperation and the use of relays have been proposed to enhance the performance of general communication systems and cognitive radio networks. For instance, relays are used to assist primary transmission to compensate for the interference from the secondary transmission in [4] and [5]. Other works discuss relaying for the secondary system during the periods of primary inactivity [6]. In [7]–[9], distributed beamforming employing a number of relays is used to improve the reliability of secondary transmissions while controlling the interference that may accrue at the primary receivers.

Another cooperative diversity technique that uses only a single best relay was first proposed in [10]. This cooperative technique has been shown to achieve diversity gain without the overhead and reduction in spectral efficiency characteristic of the concurrent operation of multiple relays. In [10]–[12], the selection of the best relay is decentralized and based on a timer initialization at each relay. Based on the channel gains of each relay to the receiver, the countdown of the timers finish at different times. Assuming that a relay declares itself once its timer is finished, the best relay can be identified as the one to first declare that its timer has run out. Although such techniques provide diversity gains as in other multiple relay cooperative schemes, synchronization errors and propagation delays between the relays cause transmission collisions as mentioned in [10], [11].

The technique for best relay selection using countdown timers was extended to cognitive radio networks (CRNs) in [13], [14] where the relays have to satisfy an interference constraint in order to coexist with the primary user. Design and analysis in [13], [14] assume that no collisions take place in the selection process. It is worth mentioning that the collisions may be more significant in a cognitive setting than in a conventional network due to the additional interference constraints.

The main contribution in this paper is that we account for transmission collisions in a relay-assisted two-hop cognitive radio system. We propose discrete countdown timers with optimized boundaries for decentralized relay selection. We provide closed form expressions for the outage probability of the secondary link using discrete timers and under both the decode-and-forward and amplify-and-forward modes of operation. The timer boundaries are then optimized to minimize the outage probability which is proportional to the collision probability between the relays.

The remainder of this paper is divided as follows. In Section II, we describe the system model used in our problem and the relaying operation decisions taken by the transmitters and relays in the system. A through analysis of the outage probability in the different scenarios of the best relay selection scheme is provided in Section III. Next, we evaluate the scheme with numerical results in Section V and finally provide some concluding remarks in Section VI.

II. SYSTEM MODEL & OPERATION

We consider an underlay cognitive network consisting of a secondary transmitter–receiver pair (ST and SR, respectively) and a relay cluster of N relays to assist communication between ST and SR, as shown in Fig. 1. It is assumed that direct link from the ST and SR suffers from severe pathloss and, hence, the SR only receives information from the ST via the relays.
The primary system is assumed to be continuously transmitting during all time slots on the channel of interest. The secondary network should constrain the interference it causes at the primary receiver (PR). Interference caused to the PR can be the result of transmission from the ST or the "best" relay, $R_i$, selected for transmission as discussed below. For the primary transmission to be decodable, the interference power at the PR should not exceed a maximum threshold, $\alpha$, which is known to all nodes in the secondary network.

We consider three sets of channels in our model: $f_i$, $h_i$ and $g_i$. $f_i$ is the channel coefficient between the relay $R_i$ and PR, $h_i$ is the channel between the ST and $R_i$, and $g_i$ is the channel between $R_i$ and the SR, where $i = 1, 2, 3, \ldots, N$. The channel between the ST and PR is denoted as $f_0$. All channels in the secondary network, i.e., $h_i$ and $g_i$, are assumed to be zero-mean complex Gaussian fading channels with variance $\sigma^2$ and to be independent of one another. The interference channels from the secondary nodes to the PR, $f_i$, are assumed to zero-mean complex Gaussian fading channels with variance $\sigma_i^2$ and to be independent of one another and of the channels in the secondary network. We adopt a slow fading channel model where the channel gains remain constant over tens of time slots. Channel estimation occurs during a fraction of the time slots by listening to dedicated symbols for channel estimation transmitted by the secondary nodes and by overhearing the automatic repeat request (ARQ) feedback packets that PR sends to its respective transmitter. Due to the broadcast nature of wireless communications, these packets can be overheard by the secondary nodes and their received signal strength can be used to estimate the channel gains to the PR.

The relays may operate in an amplify-and-forward (AF) or decode-and-forward (DF) manner. In the AF case, the relay just amplifies the signal it receives from the ST and forwards it to the SR while in DF, the relay decodes the signal it receives from the ST and forwards that signal if decoding is successful. The signal-to-noise ratio (SNR) at the SR in the AF case when it receives from $R_i$, $\gamma_i$, is given by [15]:

$$\gamma_i = \frac{|h_i|^2|g_i|^2 P_i P_r}{\sigma^2 \sigma_i^2 + |h_i|^2 P_s \sigma_s^2 + |g_i|^2 P_r \sigma_r^2}$$  \hspace{1cm} (1)

where $P_s$ is the transmit power of the ST, $P_r$ is the transmit power of the relay, $\sigma_s^2$ is the noise power at the relay, and $\sigma_r^2$ is the noise power at the SR. We assume that all relays have the same transmit power $P_r$ and noise power $\sigma_r^2$. In the DF case, if relay $R_i$ transmits to the SR after successfully decoding the signal it gets from the ST, the receive SNR at the SR is:

$$\gamma_i = \frac{P_i |g_i|^2}{\sigma_s^2}$$  \hspace{1cm} (2)

Note that the operation of the best relay selection framework requires synchronization between the different relays. Accurate synchronization schemes have been studied in [16], [17] for single relay networks and can be used to minimize the uncertainty in time alignment. In this paper, we will assume that the relays are synchronized and focus on collision probability minimization.

We discuss now the three phases of transmission within the secondary system time slot shown in Fig. 1.

**A. First Transmission Phase**:

During the first phase within each time slot, the ST may transmit to the relays. The transmission decision is based on the channel magnitude of the ST-PR link $|f_0|$. If $P_s |f_0|^2 > \alpha$, then the ST remains silent during the time slot. Alternatively if $P_s |f_0|^2 \leq \alpha$, the ST transmits.

**B. Relay Selection**:

During the relay selection phase, the best relay is agreed upon in a distributed manner. This occurs over two steps:

1) **Local Selection:**

Each relay $R_i$ decides whether to run for best relay status or to remain silent for the rest of the time slot. For the relay to qualify for transmission, it must locally satisfy some conditions, which depend on the relay network’s mode of operation.

**Conditions for the AF case:**

a. Interference at the PR should not exceed $\alpha$:

$$|f_i|^2 \leq \frac{\alpha}{P_r}$$

b. Signal received at the SR through relay $R_i$ is decodable:

$$\frac{\gamma_i}{\beta_i} \geq \gamma_{th}$$

where $\gamma_{th}$ is the minimum SNR required at SR for successful decoding of the signal from ST.

**Conditions for the DF case:**

a. Interference at the PR should not exceed $\alpha$:

$$|f_i|^2 \leq \frac{\alpha}{P_r}$$

b. The signal received at the SR from $R_i$ is decodable:

$$\gamma_i \geq \gamma_{th}$$

c. The signal received at $R_i$ from the ST is decodable:

$$\beta_i \geq \beta_{th}$$

where $\beta_i$ is the channel gain between the ST and $R_i$, $\beta_i = \frac{P_s |h_i|^2}{\sigma_s^2}$, and $\beta_{th}$ is the minimum SNR required at any relay for successful decoding of the signal from the ST.

2) **Global Selection**:

The relays selected from the previous step for satisfying their local conditions run for a global selection of the best relay. Implementation of best-relay selection is done using a discrete time-out timer. As shown in Fig. 2, the timer is characterized by the parameter $K$ and boundary values $b_1, b_2, \ldots, b_K$, which are employed at all relays. At relay $R_i$, the timer is initialized on the basis
We now study the remaining two conditions for the AF and DF modes.

A. Amplify-And-Forward:

A subset $M$ of $m$ out of the $N$ relays is selected such that the relays in $M$ satisfy the conditions described in II-B for the AF mode. Let $p_{sel}$ denote the probability that a single relay satisfy the AF selection conditions:

$$p_{sel} = Pr\{|f_1|^2 \leq \frac{\alpha}{P_s}, \gamma_i \geq \gamma_{th}\}$$  

(4)

Since $\gamma_i$ and $f_1$ are independent, the joint probability in (4) is the product of the probability of each event.

$$Pr\{|f_1|^2 \leq \frac{\alpha}{P_s}\} = 1 - \exp\left(-\frac{\alpha}{\sigma_f^2 P_s}\right)$$  

(5)

The probability that $\gamma_i$ exceeds $\gamma_{th}$ is given by [18]:

$$Pr\{\gamma_i \geq \gamma_{th}\} = \frac{2\gamma_{th}}{\sqrt{\pi} \eta_i} \exp\left[-\gamma_{th}\left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right)\right]I_1\left(\frac{2\gamma_{th}}{\sqrt{\pi} \eta_i}\right)$$  

(6)

where $I_1(.)$ is the first order modified Bessel function of the second kind.

$$\eta_1 = \frac{P_r \sigma_f^2}{\sigma_e^2} \text{ and } \eta_2 = \frac{P_r \sigma_f^2}{\sigma_e^2}$$  

(7)

The probability $p(m)$ of selecting a total of $m$ relays can therefore be written as:

$$p(m) = \binom{N}{m} p_{sel}^m (1-p_{sel})^{N-m}$$  

(8)

As mentioned in Section II, after each relay $R_i$ evaluates $\gamma_i$, a discrete time-out timer is initiated to flag out the single best relay. The probability $\Lambda_k$ of a timer getting initialized to $k$ given that relay $R_i$ is a member of $M$ is denoted by:

$$\Lambda_k = \int_{b_k}^{b_{k+1}} p\left(\gamma_i | \gamma_i \geq \gamma_{th}\right) d\gamma_i$$

$$= \frac{1}{Pr\{\gamma_i \geq \gamma_{th}\}} \left[\frac{2b_{K-k}}{\sqrt{\pi} \eta_i} e^{-b_{K-k}} \left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right) I_1\left(\frac{2b_{K-k}}{\sqrt{\pi} \eta_i}\right) - \frac{2b_{K-k+1}}{\sqrt{\pi} \eta_i} e^{-b_{K-k+1}} \left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right) I_1\left(\frac{2b_{K-k+1}}{\sqrt{\pi} \eta_i}\right)\right]$$  

(9)

A successful secondary transmission takes place when ST is capable of transmission and only one relay from the subset $M$ is initialized to $k$ while the remaining $m-1$ relays are initialized to higher values. Therefore, the probability of transmission success is given by:

$$p_c = Pr\{|f_1|^2 \leq \frac{\alpha}{P_s}\} \left(\sum_{m=1}^{N} p(m) \sum_{k=0}^{K} m \Lambda_k \left(\sum_{\nu=k+1}^{K} \Lambda_\nu\right)^{m-1}\right)$$  

(10)

B. Decode-And-Forward:

A subset $M$ of $m$ out of the $N$ relays is selected such that the relays in $M$ satisfy the conditions described in II-B for the DF mode. Let $p_{sel}$ denote the probability that a single relay satisfy the DF selection conditions:

$$p_{sel} = Pr\{|f_1|^2 \leq \frac{\beta_i}{\nu}, \beta_i \geq \beta_{th}, \gamma_i \geq \gamma_{th}\}$$  

(11)

Since $f_1$, $\beta_i = \frac{P_r |h|_i^2}{\sigma_f^2}$ and $\gamma_i = \frac{P_r |g|_i^2}{\sigma_e^2}$ for DF case are independent, the joint probability (11) is the product of the probability of each condition. The probability of the first
condition is given in (5). The probability of the second and third conditions on $\beta_i$ and $\gamma_i$ is given by:

$$\Pr\{\beta_i \geq \beta_{th}\} = e^{-\frac{\beta_{th}}{\eta_1}}, \quad \Pr\{\gamma_i \geq \gamma_{th}\} = e^{-\frac{\gamma_{th}}{\eta_2}}$$  \quad (12)$$

with $\eta_1$ and $\eta_2$ given by (7). The probability $p(m)$ of selecting $m$ relays in DF mode is given by (8) with $p_{sel}$ given by (11). Provided that relay $R_k$ is a member of $M$, the probability $\Lambda_k$ of a timer getting initialized to $k$, knowing $\gamma_i$ is:

$$\Lambda_k = \int_{b_{K-k}}^{b_{K-k+1}} p(\gamma_i | \gamma_i \geq \gamma_{th}) \, d\gamma_i = \frac{e^{-\frac{b_{K-k}}{\eta_2}} - e^{-\frac{b_{K-k+1}}{\eta_2}}}{\Pr\{\gamma_i \geq \gamma_{th}\}}$$  \quad (13)$$

Expression (13) is used to compute $p_{c}$ in (10) for the DF case. In other words, expressions for $p(m)$ and $p_{c}$ in (8) and (10), respectively, are used for both AF and DF cases. Nevertheless, both cases differ in their formulas for $p_{sel}$ and $\Lambda_k$.

IV. TIMER OPTIMIZATION

Using the equations developed in Section III, we can evaluate the $p_{out}$ of the secondary system. By inspection of equations (9) and (13), we can see that $\Lambda_k$ — and, hence, the outage probability, $p_{out}$ — are dependent on our choice of the boundary values, $b_k$. The design of the boundary values impacts the rate of collision between the relays and therefore affects the outage of the secondary system. Our goal in optimizing the timer design is to arrive at the boundary values that minimize $p_{out}$. We obtain the optimal boundary values by first deciding the timer initialization probabilities that minimize $p_{out}$. The relations between the timer initialization probabilities and the boundary values in (9) and (13) can be used to obtain the corresponding optimal boundary values.

Let $\Lambda$ be a vector of $K+1$ elements such that $\Lambda = [\Lambda_0, \Lambda_1, \Lambda_2, \ldots, \Lambda_K]$. We want to obtain the optimal vector $\Lambda$ that is a solution to the following optimization problem:

$$\min_{\Lambda} p_{out} = 1 - p_{c}$$
subject to: \quad $\sum_{k=0}^{K} \Lambda_k = 1$, \quad $\Lambda_k \geq 0 \, \forall k = 0, 1, \ldots, K$  \quad (14)$$

This problem can be shown to be nonconvex due to the nonconvexity of the objective function. We use the technique of genetic algorithms to obtain the solution [19].

Note that the solution of (14), albeit complex, is needed only when the statistics of the channel change. This is assumed to occur over relatively large time scales, i.e., many time slots.

After we get the solution $\Lambda$, we use (9) and (13) to calculate the optimal boundary values $b_k$. In case of DF scheme, the inverse of equation (13) is straightforward and is given by:

$$b_k = \left[ -\frac{\eta_2}{\gamma_{th}} \log_e \left( 1 - \Pr\{\gamma_i < \gamma_{th}\} \right) + \Pr\{\gamma_i \geq \gamma_{th}\} \sum_{j=K-k+1}^{K} \Lambda_j \right]$$  \quad (15)$$

such that $k = 1, 2, 3, \ldots, K$. For equation (9), the inverse function is not straightforward but boundary values $b_k$ can be obtained numerically.

V. NUMERICAL RESULTS

We provide here some numerical results exploring the impact of various parameters on $p_{out}$. We assume that $P_s = P_r = 2$ mW, threshold SNR values, $\beta_{th}$ and $\gamma_{th}$ are set to unity. The interference channels from secondary nodes to PR, $E[|h_i|^2] = \sigma_f^2 = 1$. Also noise parameters, $\sigma_f^2 = \sigma_n^2$ are set to unity.

Fig. 3 shows the variation of $p_{out}$ with the number of relays, $N$. This figure is for the AF and DF cases with $K = 100$, $\alpha = 10$ and $E[|h_i|^2] = E[|g_i|^2] = \sigma_f^2 = 3$ dB, $i = 1, 2, \ldots, K$. It is clear from the figure that $p_{out}$ decreases with $N$, reaches a minimum at $N = 9$ for the AF case and $N = 10$ for the DF case, and then increases very slightly. Under a collision-free best relay selection scenario, increasing $N$ enhances the diversity order of the system due to the increased number of transmission paths. However, the increase in diversity is countered by the increased level of collisions.

In Fig. 4, we show how $p_{out}$ changes with the number of levels allowed for the discrete timer. As shown in the figure, which is focused on the case with $N = 10$, $\alpha = 10$ and $\sigma_f^2 = 3$ dB, as $K$ increases, $p_{out}$ decreases because this allows for more fine partitioning of the SNR scale and reduces the probability of timers being initialized to the same value. First note that the minimum interval that can be used per level is the duration needed for hardware from reception to transmission [10]. Therefore, although increasing $K$ would reduce $p_{out}$, it would require a longer relay selection phase. The study and optimization of this trade-off is the subject of an ongoing investigation by the authors.

Fig. 5a shows the change of $p_{out}$ with $\sigma_f^2$ for $K = 100$, $\alpha = 10$ and two values of $N$, 5 and 10. The secondary outage probability decreases, sometimes very slowly, with increases in $\sigma_f^2$. As $\sigma_f^2$ is increased, the channels become more reliable and more relays participate in the best relay selection scheme. Hence, the probability of collisions increase given a fixed $K$, thereby causing the $p_{out}$ curve to flatten out.

Fig. 5b depicts the variation of $p_{out}$ for $K = 100$ with $\alpha$. The relevant figure parameters are provided in the figure’s legend. As $\alpha$ increases, $p_{out}$ is reduced as more relays can potentially transmit to the secondary destination.

To highlight the gain of using discrete optimized timers, we compare the discrete optimized timer design with the
Fig. 5: Effect of system parameters on the outage probability $P_{\text{out}}$.

(b) Secondary outage $P_{\text{out}}$ versus $\alpha$.

Continuous timer design proposed in the literature in [10], [11] for the DF case. The results are observed in Fig. 6a. The timer initialization of the continuous timer, $T_i$, is inversely proportional to the channel gain and is given by:

$$T_i = \frac{\Gamma}{\gamma_i} = \frac{K_{\gamma_{th}}}{\gamma_i}$$

where $\Gamma$ is a design constant which we set to $K_{\gamma_{th}}$ so that a timer is initialized to $K$ when $\gamma_i = \gamma_{th}$ and to a lower value when $\gamma_i > \gamma_{th}$. In the discrete setting we consider a collision event if two relays are initialized with the same timer initialization values. Assume that each timer is decreased by one after $T_d$ units of time and that the relays sense the channel for $T_d$ to know whether another relay will undergo the transmission to the secondary destination. For comparison purposes, we consider the continuous timer relaying system with the same $K = 20$ and $T_d = 1$ unit of time. If the best relay has $T_i = T_b$, a collision event is considered if the second best relay is initialized with a value less than or equal to $T_b + T_d$. In the figure, an approximation quantizer to the timer in (16) is plotted based on the aforementioned collision criterion. The figure shows that at higher SNR values, the discrete timer optimized with high channel variance in mind would dedicate several levels for the same range that the approximate quantizer would count as one. An example can be seen by comparing the number of levels for the continuous timer with that of the discrete timer optimized with $\sigma_2 = 13$ dB in the SNR range 50-60 dB where one level in the continuous timer is compensated by nine levels in the optimized discrete timer.

Fig. 6b shows the advantage of timer optimization over the use of the formula (16) for $K = 20$. An explanation for the increased $P_{\text{out}}$ for the continuous timer case is that for large $\gamma_i$ values, the reciprocal relation between $T_i$ and $\gamma_i$ will cause the initialization of two relays with high $\gamma_i$ to be very close. Hence, collisions occur more frequently and outage becomes more probable. In contrast, the optimal design of discrete timers adapt to the SNR distribution and is not just dependent on the instantaneous value $\gamma_i$. Therefore, collisions do not become as significant when $\sigma^2$ increases.

VI. CONCLUSION

In this paper, we study an underlay CRN where the secondary communication is performed through a number of relays. A best relay selection scheme based on timers is used during the relaying operation. We focus in this work on timer optimization to minimize the probability of outage in the secondary network, which may be caused when two or more relays are initialized with the same lowest value and, hence, their transmissions collide. Our current ongoing investigation of the subject focuses on the impact of the delay incurred due to the relay selection process and on the effect of erroneously sensing the clear-to-send signal by the best relay. We are also studying the cases where the ST and/or the relays are capable of adapting their transmit power.

REFERENCES