Opportunistic Spectral Access in Cooperative Cognitive Radio Networks

Wei Liang[1], Soon Xin Ng[1], Siavash Bayat[2], Yonghui Li[3] and Lajos Hanzo[1]
[1]School of ECS, University of Southampton, SO17 1BJ, United Kingdom.
[2] Electronics Research Institute, Sharif University of Technology, Iran.
[3]School of Electrical and Information Engineering, University of Sydney, NSW, 2006, Australia.
Email: [1]{wl7g09, sxn, lh}@ecs.soton.ac.uk ;[2]bayat@sharif.edu; [3]lyh@ee.usyd.edu.au

***Abstract*— A pragmatic distributed algorithm (PDA) is proposed for supporting the efficient spectral access of multiple Primary Users (PUs) and Cognitive Users (CUs) in cooperative Cognitive Radio (CR) networks. The CUs may serve as relay nodes for relaying the signal received from the PUs to their destinations, while both the PUs' and the CUs' minimum rate requirements are satisfied. The key idea of our PDA is that the PUs negotiate with the CUs concerning the specific amount of relaying and transmission time, whilst reducing the required transmission power or increasing the transmission rate of the PU. Our results show that the cooperative spectral access based on our PDA reaches an equilibrium, when it is repeated for a sufficiently long duration. These benefits are achieved, because the PUs are motivated to cooperate by the incentive of achieving a higher PU rate, whilst non-cooperation can be discouraged with the aid of a limited-duration punishment.**

***Index Terms*— Cognitive Radio network, Cooperative Communication, Spectrum Access strategy, Repeated game.**

I. Introduction

Cognitive radio (CR) was found to be able to improve the spectral efficiency by exploiting the available spectrum holes [1]–[3]. If the spectrum is not used by the Primary Users (PUs), then the Cognitive Users (CUs) have the opportunity to access it for their secondary communications based on the CR technique. According to the CR protocol, the device listens to the received signal and identifies the spectrum holes, using either time or frequency domain sensing [1]–[3]. In this context, the most common paradigms associated with CRs are the so-called underlay, overlay and interweave networks [3].

Cooperative communication is a novel communication paradigm that promises significant capacity and multiplexing gain improvements in wireless networks and it is capable of supporting the users either at an improved integrity or throughput with the advent of user cooperation [4], [5]. The benefits of employing cooperative communication in CR networks have been studied in [6]–[9].

Different PUs may operate using different licensed bands in the same geographical location. The authors of [10] considered multiple PUs and CUs. Moreover, all PUs and CUs may be considered to be selfish, hence they may only be concerned about their own benefit, pursuing their own best strategies for maximizing their own sum rate or throughput. A conventional distributed algorithm (CDA) was proposed in [11], that facilitates a joint competitive strategy of the PUs and the CUs conceived for accessing the spectral resources. Moreover, the authors of [11] conceived a non-cooperative game, which employs the CDA for efficiently representing the interaction among the competing PUs, where each PU chooses its allocation independently of the others in order to improve its own performance. Explicitly, a spectral access strategy is designed for multiple PUs and CUs, where the PUs and CUs are carefully paired for ensuring that both the PUs' and the CUs' minimum sum-rate requirements are satisfied. Each paired CU assists in relaying its paired-PU's signal in exchange for a transmission opportunity using the PU's spectrum. However, the PUs under the CDA would sometimes compete among themselves for cooperating with the same relay, which may degrade both their utility and throughput.

The financial support of the European Research Council's Advanced Fellow Grant, that of the India-UK Advanced Technology Centre, as well as that of the European Union's Seventh Framework Programme (FP7/2007-2013) under the auspices of the CONCERTO project (grant agreement no 288502) is gratefully acknowledged.

In contrast to the CDA, we proposed a protocol, which may be classified as a repeated game [12], [13] where all PUs are capable of cooperating with each other. Thus they are motivated to form a grand coalition [14], [15] for achieving an increased expected PU rate by discouraging the PUs from competing with each other for the same CU's assistance. Furthermore, the concept of a penalty/punishment is introduced [12], which is imposed only for a carefully selected finite period for the sake of discouraging non-cooperation. We proposed the PDA concept for efficient spectrum access, which is invoked for cooperation amongst the PUs for obtaining an improved performance for the PUs. Additionally, this PDA is formulated as a punishment based repeated game for attaining equilibrium.

The rest of the paper is organized as follows. The system model of the cooperative CR network considered is outlined in Section II, while our spectral access strategy is detailed in Section III. The performance of the proposed schemes is evaluated in Section IV. Finally, our conclusions are presented in Section V.

II. System Model

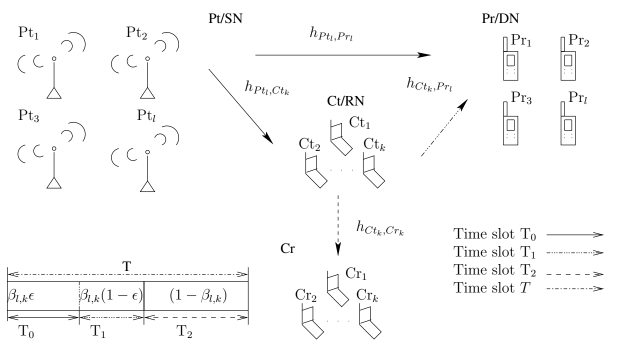


Fig. 1. The spectrum-access model for primary user and cognitive user.

In line with [11], we consider an overlay cooperative CR scheme comprising L_{PU} number of "PU transmitter (Pt) and PU receiver (Pr) pairs"[1], namely ($\{Pt_l\}_{l=1}^{L_{PU}}$,$\{Pr_l\}_{l=1}^{L_{PU}}$), with the lth pair having a rate requirement of $R_{PU_l,req}$, and with each pair occupying a unique spectral band of a constant width. In our scheme, there are L_{CU} number of "CU transmitter (Ct) and CU receiver (Cr) pairs"[2],

[1]The Pt-Pr pair constitutes the PU's transceiver. We represent the PU acting as the source node by Pt. Similarly, Pr denotes the PU which acts as the destination node PU.

[2]The Ct-Cr pair is the CU's transceiver. The notation Ct/RN represents the CU that acts as the relay node for helping the PU's transmission, while Ct denotes the CU that act as the source node (SN). Similarly, the Cr denotes the destination node of a Ct.

namely ($\{Ct_k\}_{k=1}^{L_{CU}}$,$\{Cr_k\}_{k=1}^{L_{CU}}$) pairs, with the kth pair having a requirement of $R_{CU_k,req}$, and seeking to obtain access to a spectral band occupied by a (Pt, Pr) pair. Each Pt attempts to grant spectral access to a unique (Ct, Cr) pair in exchange for the Ct cooperatively relaying the Pt's data to the corresponding Pr. Four types of matching algorithms will be described in Section III.

Fig. 1 illustrates the time period allocation of the PUs and CUs, where T is the original time period allocated for the Pt to transmit its source message to the Pr. We will refer to $\{\beta_{l,k}\}_{l=1k=1}^{L_{PU}L_{CU}}$ as the time allocation fraction, where $0 < \beta_{l,k} < 1$. When the Pt is assisted by a Ct/RN, the Pt relies on a time-fraction of $\beta_{l,k}T$ to convey the source message to the Pr and Ct/RN. More specifically, the Pt simultaneously transmits its message to Pr and Ct/RN during the $T_0 = \beta_{l,k}T\epsilon_{l,k}$ time-period, where $0 < \epsilon_{l,k} < 1$. Additionally, the Ct/RN cooperatively relays the Pt's signal to Pr in the subsequent $T_1 = \beta_{l,k}(1-\epsilon_{l,k})T$ time-periods. Then the Pr applies maximum ratio combining for detecting the signal received from the Pt during the first T_0 time period, and the signal received from the Ct/RN in the subsequent T_1 time periods. After the PU has ceased its transmission, the system will allow the CUs to transmit their informations to the other CUs by using the remaining time period of $T_2 = (1-\beta_{l,k})T$ for their own communications. In other words, a Ct/RN assists in saving some of the transmission powers of the Pt due to the reduction of the transmission period of PU from T to T_0. The Pt transmits during T_0, while the Ct/RN forwards the source message during T_1 and the Ct/SN can broadcast its message to other CUs during the time period T_2. Let us assume that the transmission power per unit frequency transmitted from the Pt is P_S watts/Hz and the target transmission rate is R_{PU} bits/s.

During the first time slot (TS) T_0, the PU broadcasts the source message x to both the Pr and to the Ct/RN. The signal received at the Pr is $y_{Pt_l,Pr_l} = \sqrt{P_S}h_{Pt_l,Pr_l}x + n_{Pt_l,Pr_l}$, and the signal received at the Ct/RN is: $y_{Pt_l,Ct_k} = \sqrt{P_S}h_{Pt_l,Ct_k}x + n_{Pt_l,Ct_k}$. During the second TS T_1 the Ct/RN would forward the source message to the Pr using the transmission power of P_{CR} watts/Hz. Similarly, the signal received by the Pr under the AAF protocol via the RD link may be expressed as: $y_{Ct_k,Pr_l} = \omega_A\sqrt{P_{CR}}h_{Ct_k,Pr_l}y_{Pt_l,Ct_k} + n_{Ct_k,Pr_l}$, where $\omega_A = \frac{1}{\sqrt{P_S|h_{Pt_l,Ct_k}|^2+N_0}}$ [16] is the amplification factor, while n_{Pt_l,Pr_l}, n_{Pt_l,Ct_k} and n_{Ct_k,Pr_l} are the Gaussian noise vectors, which have a zero mean and a noise variance of $N_0/2$ per dimension. The channel gain terms h_{Pt_l,Pr_l}, h_{Pt_l,Ct_k} and h_{Ct_k,Pr_l} are assumed to be Rayleigh distributed, obeying the complex-valued Gaussian distribution of $\mathcal{CN}(0,1)$. In our system, the path loss is included in the channel gain term. In our scheme, both the Pt and the Ct/RN utilize the same bandwidth. The achievable instantaneous rate of the lth PU when employing the kth CU at a given $\beta_{l,k}$ may be represented as:

$$R_{l,k}^{PU}(\beta_{l,k}) = \frac{\beta_{l,k}T}{2}\log_2\left[1 + \frac{\gamma_{PU}|h_{Pt_l,Pr_l}|^2}{d_{Pt_l,Pr_l}^{\alpha}} + f_{Pt,Ct,Pr}\right] \quad (1)$$

where we have

$$f_{Pt,Ct,Pr} = \frac{\gamma_{PU}\gamma_{CU}|h_{Pt_l,Ct_k}|^2|h_{Ct_k,Pr_l}|^2}{\gamma_{PU}|h_{Pt_l,Ct_k}|^2 d_{Ct_l,Pr_l}^{\alpha} + \gamma_{CU}|h_{Ct_k,Pr_l}|^2 d_{Pt_l,Ct_l}^{\alpha} + d_{Pt_l,Ct_l}^{\alpha}d_{Ct_l,Pr_l}^{\alpha}} . \quad (2)$$

The factor $\frac{1}{2}$ in Eq. (1) is due to the time fraction $\epsilon_{l,k} = \frac{1}{2}$, when we have $T_0 = T_1$, where the Pt utilizes the first TS T_0 and the Ct/RN uses the second TS T_1 to transmit the PU's signals. Note that the transmit SNR[3] of the PU is $\gamma_{PU} = \frac{P_s}{N_0}$ and that of the CU is $\gamma_{CU} = \frac{P_{CR}}{N_0}$. Moreover, in the non-cooperative scenario, the achievable sum rate of the direct link between Pt_l and Pr_l is given by:

$$C_{PU}^{*} = T\log_2\left[1 + \frac{\gamma_{PU}|h_{Pt_l,Pr_l}|^2}{d_{Pt_l,Pr_l}^{\alpha}}\right] , \quad (3)$$

while the minimum rate requirement of the PU is given by $R_{l,req}^{PU} = C_{PU}^{*}$.

The achievable transmission rate of the kth CU when assisting the lth PU at a given $\beta_{l,k}$ is formulated as:

$$R_{l,k}^{CU}(\beta_{l,k}) = (1-\beta_{l,k})T\log_2\left[1 + \gamma_{CU}|h_{Ct_k,Cr_k}^{(l)}|^2\right] , \quad (4)$$

where the channel $h_{Ct_k,Cr_k}^{(l)}$ depends on the frequency band provided by Ct_l, while the pathloss is $\varrho = 1/d_{ab}^{\alpha}$ [16] and d_{ab} is the geometrical distance between node a and node b. The path-loss exponent is given by $\alpha = 4$. Moreover, we assume that Pt and Pr are located at the opposite sides of a square at a normalized distance of two, thus $d_{Pt_l,Pr_l} = 2.0$ as considered in [11]. Additionally, the Cts/RN and Crs are assumed to be randomly located within an internal square having an edge-length of one, hence we have $(0 < d_{Ct_k,Cr_k} < \sqrt{2})$.

III. Spectrum Access Algorithm

In this section, we briefly highlight the algorithm invoked for determining the spectral access for each (Pt, Pr) and (Ct, Cr) pairs.

A. Preference Lists

Before any offer is made to the CUs, the PUs construct a preferred list of CUs, which can satisfy the PU's rate requirement. Thus, the preference list for Pt_l is given by:

$$PULIST_l = \left\{(Ct_{\kappa(k)}, Cr_{\kappa(k)})\right\}_{k=1}^{L_{CU}} , \quad (5)$$

where the function $\kappa(k)$ satisfies the following conditions:

$$R_{l,\kappa(k)}^{PU}(\beta_{l,\kappa(k)}) > R_{l,req}^{PU},\ k \in (1,\ldots,L_{CU}) . \quad (6)$$

We consider the indexes of users inside the preference lists not the system rate. Additionally, we have assumed that the first $\mathrm{Ct}_\kappa(k)$ at the top of the $PULIST_l$ provides the highest rate $R_{l,\kappa(k)}^{PU}(\beta_{l,\kappa(k)})$. Similarly, each CU also has its preferred PU list, and if it transmits in the spectral band occupied by the preferred PUs then its achievable transmission rate is higher than its minimum sum-rate requirement, $R_{k,req}^{CU}$. Thus, the preference list for Ct_k is given by:

$$CULIST_k = \left\{(Pt_{\iota(l)}, Pr_{\iota(l)})\right\}_{l=1}^{L_{PU}} , \quad (7)$$

where the function $\iota(l)$ satisfies the following conditions:

$$R_{\iota(l),k}^{CU}(\beta_{\iota(l),k}) > R_{k,req}^{CU},\ l \in (1,\ldots,L_{PU}) . \quad (8)$$

Again, the ordering of the CULIST_k also range from the highest to the lowest.

[3]The definition of the transmit SNR is unconventional, because it relates the transmit-power to the noise-power at the receiver, which are quantities measured at different locations. Nonetheless, this convenient definition simplifies our discussion, as proposed in [17].

B. Conventional Distributed Algorithm (CDA)

The key idea of the CDA is detailed in [11], where each (Pt, Pr) pair trades with a particular (Ct, Cr) pair for the sake of attaining mutual benefits in the context of cooperative relaying. More specifically, the CDA constitutes a non-cooperative scheme, where none of the PUs cooperates. Instead, they compete with each other, with the selfish objective of maximizing their own rate. Let us denote the average rate of PU_l in CDA as:

$$r_l^S = E\left[R_l^S\right] \ , \tag{9}$$

where $E[.]$ is the expected value of [.], the superscript S indicates the selfish nature of the CDA and R_l^S is the instantaneous rate of PU_l during a particular transmission.

C. The Proposed Pragmatic Distributed Algorithm (PDA)

Our PDA is a cooperative scheme, where all PUs forms a grand coalition [15]. A *game unit* is constituted by L_{PU} rounds and each round has L_{PU} transmissions, where the PUs take turns to select the best available CU according to a round-robin type priority access list. The priority access list of the ith round is given by:

$$ALIST_i \quad = \quad \{Pt_i, Pt_{i\oplus 1}, \ldots, Pt_l, \ldots, Pt_{i\oplus(L_{PU}-1)}\} \ , \tag{10}$$

where $i = \{1, 2, \ldots, L_{PU}\}$ and the subscript of Pt_l for the jth transmission ($j = \{1, 2, \ldots, L_{PU}\}$) in the ith round is based on the modulo-L_{PU} summation:

$$l = i \oplus (j-1) = (i + (j-1)) \mod L_{PU} \ . \tag{11}$$

The first Pt in the $ALIST_i$ has the first priority to select its best CU. Then the second Pt in the list selects the best available CU from the remaining set of CUs, and so on. During the next round, the first Pt in $ALIST_i$ will become the second Pt in $ALIST_{i+1}$, while the last Pt in $ALIST_i$ is now the first Pt in $ALIST_{i+1}$ according to the round-robin scheduling. Hence, after L_{PU} rounds each PU is guaranteed to have access to $\min\{L_{PU}, L_{CU}\}$ CUs amongst the top CUs in its PULIST[4], but is has no access to any CUs for the remaining $(L_{PU} - L_{CU})$ transmissions. In this way, the PUs give up any futile competition and cooperatively take turns, one at a time, to access the available CUs, which is expected to yield the most benefits for themselves. If none of the CUs in the current list may be satisfied, then only this specific Pt will update its TS allocation and then produces a new preference list. The proposed PDA has a low complexity because it does not require any exchange of information amongst the PUs, such as their rates. The first access list $ALIST_1$ can be distributed at the beginning of the game by a PU acting as a cluster-head and the remaining lists can be computed locally by each PU. Then the PDA can be repeated automatically as many times as needed. Hence, each PU is guaranteed to have access to its top CU at least $1/L_{PU}$ times[5] on average. Note that the maximum TS allocation representing the maximum transmission period for the PU can be derived from Eq. (4) as:

$$\beta_{l,k}^{\max} \quad = \quad 1 - \frac{R_{k,req}^{CU}}{\log_2\left[1 + \gamma_{CU}|h_{Ct_k,Cr_k}^{(l)}|^2\right]} \ , \tag{12}$$

while the minimum TS allocation can be computed from Eq. (1) as:

$$\beta_{l,k}^{\min} \quad = \quad \frac{2R_{l,req}^{PU}}{T\log_2[1 + \frac{\gamma_{PU}|h_{Pt_l,Pr_l}|^2}{d_{Pt_l,Pr_l}^{\alpha}} + f_{Pt,Ct,Pr}]} \ . \tag{13}$$

The specific details of the algorithm can be summarized as follows:

[4]Provided that the rate requirements in Eq. (6) and Eq. (8) are satisfied.

[5]The PU would have access to its top CU in all L_{PU} rounds, if this CU is not sought after by other PUs.

1) Initialization:
   a) Set up the first priority list $ALIST_1 = \{Pt_1, Pt_2, \ldots, Pt_{L_{PU}}\}$ and broadcast it to all PUs.
   b) Each Pt computes the remaining priority lists $ALIST_i$ for $i = \{2, 3, \ldots, L_{PU}\}$, based on the round-robin method given in Eq. (10).
   c) Compute $\beta_{l,k}^{\min}$ and $\beta_{l,k}^{\max}$.
   d) Set $i = 1$ for the first round.
2) Do the matching for the ith round:
   a) Set the initial TS allocations to β_{init}, and set the step size of TS increment to τ.
   b) Construct $PULIST_l$ according to Section III-A based on β_{init}, where $l = \{1, \ldots, L_{PU}\}$.
   c) Construct $CU_{NOTMATCH} = \{Ct_1, \ldots, Ct_{L_{CU}}\}_{l=1}^{L_{CU}}$ to list all participating Cts.
   d) Set $j = 1$ for the first transmission.
   e) Do the matching for the jth transmission:
      i) Find the corresponding Pt_l for transmission, where $l = (i + (j-1)) \mod L_{PU}$.
      ii) Pt_l selects the best available Ct_k from $CU_{NOTMATCH}$ based on $PULIST_l$:
         - Pt_l offers $\beta_{l,k}$ to Ct_k.
         - If $\beta_{l,k}^{\min} \leq \beta_{l,k} \leq \beta_{l,k}^{\max}$, then ($Pt_l$, Pr_l) and (Ct_k, Cr_k) are matched. Remove Ct_k from $CU_{NOTMATCH}$. If $CU_{NOTMATCH} \in \emptyset$ go to Step 3, else go to Step 2f.
         - Otherwise, reduce the TS allocation to $\beta_{l,k} = \beta_{l,k} - \tau$ and update $PULIST_l$.
         - If $PULIST_l$ is empty then Pt_l is left unmatched and proceed to Step 2f.
         - Otherwise, find another match at Step 2(e)ii.
   f) Set $j = j + 1$ and go to Step 2e for the next transmission, until $j = L_{PU}$.
3) Set $i = i + 1$ and go to Step 2 for the next round, until $i = L_{PU}$.
4) Terminate the game or repeat the game from Step 1 until no more transmission is needed.

The rate of PU_l averaged over L_{PU} transmissions in the ith round can be computed as:

$$R_l^C = \frac{1}{L_{PU}} \sum_{i=1}^{L} R_{l,\kappa(i)}^{PU}(\beta_{l,\kappa(i)}) \ , \tag{14}$$

where the superscript C signifies the cooperative nature in PDA, $L = \min\{L_{PU}, L_{CU}\}$ and $\kappa(i)$ is the index of the best available Ct which satisfies the rate conditions of Eq. (6) and Eq. (8) during the ith round, while $R_{l,\kappa(i)}^{PU}(\beta_{l,\kappa(i)}) = 0$ if $\kappa(i) \in \emptyset$. Hence, the average rate of PU_l after many repetitions is given by:

$$r_l^C = E\left[R_l^C\right] \ . \tag{15}$$

The proposed PDA does not require any exchange of the PU's rate information and we assume a practical time-varying wireless channel, which may change for each transmission round and the users are also allowed to move. According to the law of large numbers, once the PDA has been repeated a sufficiently high number of times, all PUs will achieve the same average PU rate due to having random channel conditions[6].

D. Centralized Algorithm (CA) and Random Algorithm (RA)

In the centralized algorithm (CA) [11], we consider all possible matching of the (Pt, Pr) and (Ct, Cr) pairs, and then select that particular matched pair, which has the maximum sum rate. By definition, in our cooperative CR scheme we give higher priority to the PUs. In particular, we focus our attention on maximizing the PUs' utility. For the RA, each Pt will make an offer β to a Ct, which

[6]User mobility is considered in the channel.

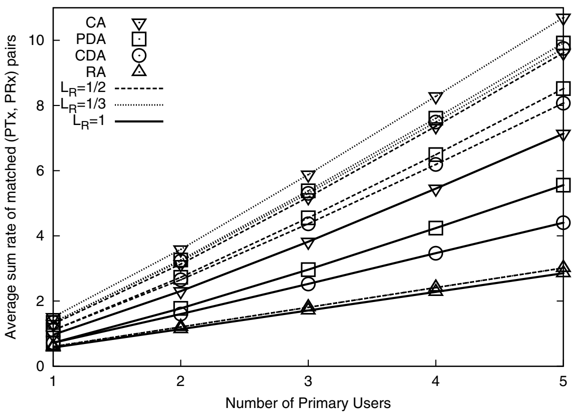
is randomly selected from its preference list. More specifically, we followed the approach adopted in [11], where this β value is fixed for all PU and SU pairs and it was chosen experimentally for the sake of maximizing the average PU sum-rate.

Note that the CA and the RA are used to benchmarking both the CDA and the proposed PDA and represent the two extremes in terms of their overhead and the complexity imposed.

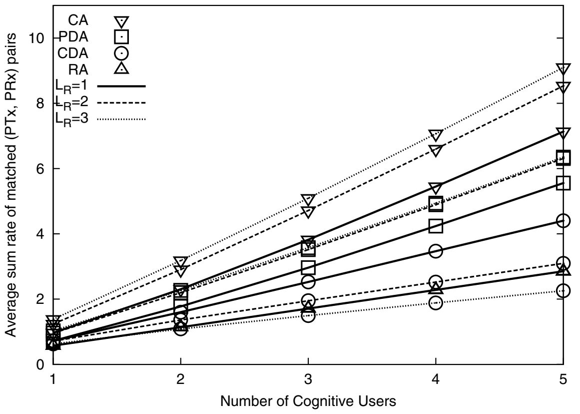
IV. Performance Results

A. Performance of the matching algorithm aided cooperative CR network

Fig. 2 shows the average total sum-rate of the matched (Pt, Pr) for the proposed cooperative CR scheme. In our evaluation of Eq. (1), Eq. (3) and Eq. (4), we assumed that the transmit SNRs of all CUs are equal, yielding $\gamma_{CU_1} = \cdots = \gamma_{CU_l} = \gamma_{CU} = 25$ dB. We also assumed that the SNR of all PUs are $\gamma_{PU_1} = \gamma_{PU_2} = 10$ dB. We have considered the CDA, the PDA, the CA and the RA in Fig. 2. The CU's rate is kept to the minimum of $R^{CU}_{k,req} = 2$ bps and the fraction L_R is given by: $L_R = L_{PU}/L_{CU}$.



(a) R^{PU}_{match} versus L_{PU}



(b) R^{PU}_{match} versus L_{CU}

Fig. 2. Performance of the **CCMC** aided **AAF based** cooperative CR **benchmark [11]** scheme communicating over quasi-static Rayleigh fading channel. The "CA", "PDA", "CDA" and "RA" techniques were detailed in Section III. The initial TS allocation is $\beta_{init} = 0.99$, the step size of TS is $\tau = 0.05$ and $L_R = L_{PU}/L_{CU}$.

Explicitly, the number of PUs and CUs is identical for $L_R = 1$, while the number of CUs is higher than that of the PUs for $L_R < 1$, and vice versa. The CA achieves the highest average total sum-rate among these four algorithms, while the RA achieves the lowest sum rate in Fig. 2(a). It is observed in Fig. 2(a) that our PDA achieves a higher sum-rate than that of the CDA at the same L_{PU}. The PDA consistently attains a higher rate than the CDA for the scenario, where the number of CU is higher than that of the PUs.

As shown in Fig. 2(b), we observe that the rate of PU$_l$ operating under our PDA is lower than that of the CA, but much higher than that of the CDA, when the number of PUs is higher than that of the CUs. Furthermore, when we have $L_R = 3$, the CDA scheme performs slightly worse than the RA scheme, which is a consequence of the competition loss encountered. Moreover, the performance curves of RA are overlapped, since increasing the number of PUs or CUs does not influence its random selection. Additionally, the average individual PU rate was found to be exactly a fraction of $\frac{1}{L_{PU}}$ of the total PU sum rate for all PUs. Thus the trend of the individual PU rate is the same as that of the total PU sum rate. Hence, the proposed PDA outperforms the CDA in terms of both the total PU sum rate and the individual PU rate, especially when the number of PUs is higher than that of the CUs, i.e. for $L_R > 1$.

B. Stability of the PDA

The CDA was shown in [11] to create a stable matching, which exhibits a *competitive equilibrium*, when all the PUs are non-cooperative. If we consider a 'single-shot' game, where each PU only cares about its current payoff, then no individual PU would have the incentive to deviate from the CDA strategy. Hence, the CDA may be deemed to be a strategy that arrives at an equilibrium for the one-shot non-cooperative game, having an expected one-shot payoff of R^S_l given in Eq. (9). However, spectrum sharing between PUs and CUs may last for a long period of time, which may be viewed as a game repeated for many rounds, in which the PUs can cooperate based on their individual reputation and their mutual trust. More specifically, the proposed PDA is capable of guaranteeing a higher individual average PU rate compared to that of the CDA due to the avoidance of competition among the PUs. As shown in Section IV, the PDA outperforms the CDA, especially when the number of CUs is lower than that of the PUs. However, a PU using the PDA may be tempted to abandon cooperation for the sake of gaining a higher instantaneous rate. Hence, we considered a penalty/punishment based repeated game [12], [13], where the PUs (players) have incentives to cooperate for the sake of achieving a higher expected payoff (average PU rate), while any non-cooperative behaviour can be avoided by appropriate punishment over a carefully selected limited period. Although the PDA may not converge to a stable equilibrium in a single-shot game, it does converge to an equilibrium in the repeated game enforced by the threat of punishment. If any of the PUs opts out of cooperation in the PDA, all PUs would revert to the non-cooperative CDA for a period of sufficiently long duration. This punishment would discourage opting out and would help to maintain cooperation.

More explicitly, the payoff of PU$_l$ in a repeated game is defined as the sum of payoffs of PU$_l$ discounted over time according to [12]:

$$U_l = (1-\delta)\sum_{i=1}^{\infty} \delta^i R^C_l[i] \ , \quad (16)$$

where δ ($0 < \delta < 1$) is the discount factor and $R^C_l[i]$ is the average rate of PU$_l$ defined in Eq. (14) for the ith round. When we have $\delta \to 1$, the PU is more patient and hence any future reward is weighed identically to the current payoff. Hence, the PU will constrain its current behaviour in the interest of maintaining a good reputation. Let us denote the discounted payoff for cooperation in the PDA as U^C_l and that for opting out of cooperation (deviation) in the CDA as U^D_l. Then the following proposition[7] suggests that both U^C_l and

[7]The proofs of all propositions are given in the Appendix.

U_l^D would converge to their means. Hence, it is better to maintain cooperation for each PU's benefit, as long as, we have $r_l^C > r_l^S$.

Proposition 1: As we have $\delta \rightarrow 1$, the instantaneous payoff U_l^C would converge to the expected payoff r_l^C while the current payoff U_l^D would converge to the averaged payoff r_l^S.

On the other hand, imposing an infinite-duration punishment is not efficient for all PUs, because all of them would be punished and would only result in a reduced PU rate of r_l^S. Explicitly, a *limited-duration punishment* [12] is a more efficient way of preventing non-cooperation, as long as the punishment is long enough to negate the one-time non-cooperation gain. If any PU deviates from cooperation in the PDA, then all PUs would revert to the non-cooperative CDA for a punishment period of T_p instances. Next, we show in the following proposition that the limited-duration punishment based cooperation in PDA also has a perfect subgame equilibrium [12, Section 14.8], which ensures optimality for subgames starting from any round of the entire repeated game.

Proposition 2: Provided that $r_l^C > r_l^S$ for all l, $l \in \{1, 2, \ldots, L_{PU}\}$, we have $\tilde{\delta} < 1$, so that for a sufficiently large discount factor $\delta > \tilde{\delta}$, the game has a perfect subgame equilibrium with a discounted utility of r_l^C, provided that all players are governed by the limited-punishment strategy.

Additionally, the punishment period can be shown to be bounded by:

$$T_p \quad > \quad \max_l \frac{\tilde{R}_l^D - R_{l,req}^{PU}}{r_l^C - r_l^S} \ , \tag{17}$$

where $\tilde{R}_l^D$ is the one-time deviation gain. In other words, PU_l would not deviate from the cooperative strategy as long as T_p satisfies Eq. (17). Hence, the repeated game based on the PDA is capable of providing higher individual PU rates as well as a higher PU sum-rate on average. Therefore it is an attractive and stable game, even when no information is available about PU rates.

V. Conclusions

We have considered four algorithms conceived for spectral access in our cooperative CR scheme. We demonstrated that the proposed PDA is capable of converging to an equilibrium in the repeated game, which is a benefit of imposing a carefully chosen limited punishment period. Our numerical analysis revealed that the proposed PDA has a low complexity and achieved a better performance than the CDA based benchmark scheme, especially when the number of CUs is lower than that of the PUs.

Appendix

Appendix A: Proof of Proposition 1

Proof: The discounted payoff in Eq. (16) can be shown to be asymptotically equivalent to the average of the one-time payoffs, when δ approaches unity as follows:

$$\begin{aligned} \lim_{\delta\rightarrow 1} U_l &= \lim_{\delta\rightarrow 1}\lim_{N\rightarrow\infty} \frac{1-\delta}{1-\delta^N}\sum_{i=1}^{N}\delta^i R_l^C[i] \\ &= \lim_{N\rightarrow\infty}\sum_{i=1}^{N}\left(\lim_{\delta\rightarrow 1}\frac{\delta^i-\delta^{i+1}}{1-\delta^N}\right)R_l^C[i] \\ &= \lim_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}R_l^C[i] = r_l^C \ , \end{aligned} \tag{18}$$

where the last equality is derived from L'Hôpital's rule. Furthermore, if PU_l opts out of cooperation at instant T^*, then its payoff would be given by $\{R_l[i], i = 0, 1, \ldots, T^* - 1\}$, which are i.i.d. random variables due to i.i.d. random channel variations and its mean is given by r_l^C of Eq. (15). The payoff after the deviation is given by $\{R_l[i], i = T^* + 1, T^* + 2, \ldots\}$ which is an i.i.d. random variable with the mean r_l^S of Eq. (9). The payoff after abandoning cooperation converges to its mean, $U_l^D \rightarrow r_l^S$, due to the law of large numbers. Hence, abandoning cooperation only benefits PU_l at instant T^*. Similarly, the payoff $\{R_l[i], i = 0, 1, \ldots\}$ would converge to its mean, $U_l^C \rightarrow r_l^C$ if abandoning cooperation never happens.

Appendix B: Proof of Proposition 2

Proof: Since the CDA has an equilibrium for the one-shot game, all PUs would not disagree to adopt the CDA strategy for the punishment stage. Since the CDA would result in r_l^S, which is lower than r_l^C, the threat of using the CDA strategy as the punishment after any non-cooperation would result in only a one-time gain for the 'defector', which can be readily negated by limited-duration punishment, when we have $\delta \rightarrow 1$. Hence, all PUs are motivated to adopt the cooperative strategy of the PDA throughout the game, which is also the optimal behaviour for any subgame.

References

[1] F. K. Jondral, "Cognitive radio: A communications engineering view," *IEEE Wireless Communications*, vol. 14, pp. 28 –33, August 2007.

[2] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, pp. 13 –18, Aug 1999.

[3] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, pp. 894 –914, May 2009.

[4] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, pp. 878 –893, May 2009.

[5] L. Hanzo, J. Akhtman, M. Jiang and L. Wang, *MIMO-OFDM for LTE, WIFI and WIMAX: coherent verus no-coherent and cooperative Turbo-Tranceivers*. John wiley-IEEE Press, 2011.

[6] D. Li, Y. Xu, X. Wang, and M. Guizani, "Coalitional game theoretic approach for secondary spectrum access in cooperative cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 844–856, March 2011.

[7] Y. Han, A. Pandharipande, and S. Ting, "Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4945 –4950, October 2009.

[8] J. Feng, R. Zhang, L. Hanzo, and S. X. Ng, "Cooperative Medium Access Control Based on Spectrum Leasing," *IEEE Transactions on Vehicular Technology*, vol. 63, pp. 297–307, Jan 2014.

[9] W. Liang, S. X. Ng, and L. Hanzo, "Cooperative communication between cognitive and primary users," *IET Communications*, vol. 7, pp. 1982–1992, Nov 2013.

[10] P. Lin, J. Jia, Q. Zhang, and M. Hamdi, "Dynamic spectrum sharing with multiple primary and secondary users," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 1756 –1765, May 2011.

[11] S. Bayat, R. Louie, Y. Li, and B. Vucetic, "Cognitive radio relay networks with multiple primary and secondary users: Distributed stable matching algorithms for spectrum access," in *2011 IEEE International Conference on Communications (ICC)*, pp. 1 –6, June 2011.

[12] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.

[13] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Repeated open spectrum sharing game with cheat-proof strategies," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 1922–1933, April 2009.

[14] R. B. Myerson, *Game Theory Analysis of Conflict*. Harvard University Press, 1997.

[15] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, pp. 77–97, September 2009.

[16] K. J. R. Liu, A. K. Sadek, W. Su and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, New York, 2009.

[17] B. Zhao and M. C. Valenti, "Distributed turbo coded diversity for relay channel," *Electronics Letters*, vol. 39, pp. 786 – 787, May 2003.