Adaptive Channel Selection among Autonomous Cognitive Radios with Imperfect Private Monitoring

Zaheer Khan(✉) and Janne Lehtomäki

CWC, University of Oulu, Oulu, Finland
zaheer@ee.oulu.fi

Abstract. We analyze the problem of autonomous cognitive radios (CRs) competing for multiple potentially available channels that may offer different rewards due to their non-homogeneity. The non-homogeneity in channels may lead to payoff distribution conflict among CRs, as each CR would prefer to select the more desirable channels. In our model, CRs are not able to observe the channel selections of other competing CRs. Rather, they get an imperfect signal from which the channel selections must be inferred. We study an adaptive *win-shift lose-randomize (WSLR)* strategy that (without centralized coordination) enables the CRs to autonomously reach an efficient and fair payoff distribution outcome. We study the autonomous channel selection problem under different primary user (PU) occupancy models; analyze the proposed strategy under imperfect signals. We also investigate the impact of deviations by a selfish CR on the performance of the proposed strategy.

Keywords: Adaptation · Autonomous cognitive radios · Heterogeneous channels · Opportunistic spectrum access

1 Introduction

Both the FCC and a recent EU report have recommended the adoption of spectrum sharing technologies, including cognitive radio (CR), to address the rising demand for high-bandwidth wireless service [1,2]. Regulatory bodies are currently defining how cognitive systems with dynamic spectrum access capabilities will be allowed to operate [3].

CR wireless systems are a collection of wireless network entities that are able to adapt intelligently to the environment through observation, exploration and learning. In sensing-based OSA, cognitive radios (CRs) monitor the environment to reliably detect the primary user (PU) signals and operate whenever the band is empty. In practice, detection of PUs may rely on a combination of sensing and the use of geolocation spectrum occupancy databases [4].

When autonomous CRs have to search multiple potentially available channels for spectrum opportunities, they face competition from one another to access a channel. The end result of this competition is reduced CR throughput due to collisions among CRs that transmit simultaneously in the same channel.

D. Noguet et al. (Eds.): CROWNCOM 2016, LNICST 172, pp. 594–604, 2016.
DOI: 10.1007/978-3-319-40352-6_49

Moreover, further payoff distribution conflict among multiple CRs may arise when potentially available channels offer different rewards due to their non-homogeneity.

The following example illustrates this payoff distribution conflict. Consider a communication system with two autonomous CRs and two potentially available channels. Suppose that the reward that a CR receives from the use of a given channel is proportional to the probability that the channel is free from PU activity. Consider the case where channel A is available with greater probability than channel B. When the two CRs choose their actions autonomously and without coordination, then if a particular channel is simultaneously sensed free by the two CRs and both of them decide to transmit on the channel, a collision occurs. To maximize the total system reward would require the CRs to autonomously arrive at orthogonal channel utilizations, where one CR opportunistically utilizes channel A and the other CR opportunistically utilizes channel B. Clearly, there is a source of conflict for the CRs over such an orthogonal channel utilization, as each CR prefers the orthogonal outcome in which it selects channel A and the other CR selects channel B.

In this paper, the question we seek to answer is how CRs can autonomously arrive at an outcome that maximizes the total average reward (the total average number of successful transmissions) in the distributed CR network in a way that also minimizes the payoff distribution conflict among autonomous CRs.

In this paper, we evaluate an adaptive WSLR strategy that maximizes the total average reward in the network by leading to reduced likelihood of collisions among CRs. The proposed strategy also leads the autonomous CRs to engage in intertemporal sharing of the rewards from cooperation. The concept of fairness we focus on is envy-freeness [5]. Using simulation results we also compare the performance of our proposed strategy against other existing strategies. We study channel selection among autonomous CRs under imperfect private monitoring which implies that CRs are not able to directly see the actions of their opponents. Rather, they get an imperfect signal from which the channel selection must be inferred. We explore the impact of false alarms, channel errors, and co-channel interference tolerance on the behavior of autonomous CRs and on the performance of the proposed adaptive strategy. By co-channel interference tolerance, we mean that there is a small but non-zero probability that, even when two CRs transmit simultaneously on the same channel, one of the transmissions is correctly decoded by its intended receiver. We also evaluate the effect of varying the number of channels that a CR can sense on the performance of the proposed scheme. Moreover, we also show that the performance of the proposed strategy is not strongly affected by different primary user (PU) occupancy models.

The works in [6–8] have also studied the problem of multichannel cognitive medium access and proposed learning and allocation strategies for distributed CRs. However, the works in [6–8] assumed that each CR cooperatively follows the same strategy (protocol). Unfortunately, in the presence of non-homogeneous spectrum resources this assumption is not valid, as non-homogeneity of spectrum resources may induce some CRs to deviate from the protocol to maximize their own usage at the expense of the total CR system throughput. It is useful to study these scenarios as a model in which competition and conflict among autonomous CRs searching multiple channels for spectrum opportunities is analyzed. In this paper, we build on our previous research presented in [9]. However, different from this work, in [9] the CRs are assumed to have perfect monitoring, i.e., they get a perfect signal from which the action of other CRs must be inferred. Different from our work in [10], in this paper we evaluate the impact of correlation in channel occupancy by a PU in consecutive time slots.

The paper is organized as follows. The system model is presented in Sect. 2, while in Sect. 3 we present, analyze and compare the WSLR strategy to related strategies proposed in other works. Finally, Sect. 4 summarizes our main conclusions.

2 System Model

We examine a multichannel CR network in which N autonomous CRs have M potentially available channels. Let $\mathcal{M} = \{1, 2, ..., M\}$ represent the set of (potentially available) channels and $\mathcal{N} = \{1, 2, ..., N\}$ represent the set of autonomous CRs. We investigate the proposed method under: (1) The (i.i.d.) model of PU channel occupancy (also adopted by [6]), in which for each channel, the PU activity in a time slot is independent of the PU activity in other time slots and is also independent of the PU activity in other channels; (2) The second model which considers correlation in channel occupancy by a PU in consecutive time slots. In this model, the state of each channel is described by a two-state Markov chain, with α_i indicating the transition probability for the ith channel from PU-occupied to PU-free and β_i indicating the transition probability from PU-free to PU-occupied. This PU activity model is also adopted in [11]. We show (see numerical results in Sect. 3.2, Table 1) that the proposed method is not strongly affected by the stochastic model of the PU behavior. In this work, for simplicity, we utilize the (i.i.d.) model of PU channel occupancy for theoretical analysis. The primary user duty cycle statistics are known to the autonomous CRs. In practice, the autonomous CRs may obtain the primary user duty cycle statistics through the use of geolocation databases [4]. Each CR can sense only one channel at a time and, due to hardware constraints, at any given time each CR can either sense or transmit, but not both. To protect transmissions by the incumbent, the detection probability ($P_{d,i}$) of an autonomous CR i is fixed at a desired target value, $P_{d,i} = P_d$, for all $i \in \mathcal{N}$. In practice, P_d is required to be close to 1. In the literature this is defined as the constant detection rate (CDR) requirement [12]. For a fixed target detection probability, the false alarm probability of a CR is a variable. In this paper, we consider the effect of varying probabilities of false alarm.

Different reward values may be associated with potentially available channels. The difference in reward values of potentially available channels can be due to a variety of reasons, such as channels with different bandwidth and channels with different availability for opportunistic use. In sensing-based OSA, when multiple channels are potentially available for transmissions, time-slotted multiple access is often adopted [6,13]. In such systems, the primary users and CRs are both

Table 1. Total average reward per time slot in the CR network and highest average envy ratio between a pair of CRs in the network as a function of $N = M$ for different strategies. With and without false alarms.

$P_{fa} = 0,$	$N = M = 4$		$N = M = 6$		$N = M = 8$	
	$\sum_{i=1}^{N} G_i$	Υ	$\sum_{i=1}^{N} G_i$	Υ	$\sum_{i=1}^{N} G_i$	Υ
$rand - C$	3.1803	$\frac{0.8902}{0.6804} = 1.3$	4.37	$\frac{0.8902}{0.6804} = 1.3$	5.3997	$\frac{0.8998}{0.4999} = 1.8$
$WSLR$ i.i.d	3.1856	$\frac{0.7965}{0.7962} = 1$	4.3374	$\frac{0.7226}{0.7221} = 1$	5.3910	$\frac{0.6617}{0.6612} = 1$
$WSLR$ Markov	3.2	$\frac{0.807}{0.804} = 1$	4.31	$\frac{0.7186}{0.7140} = 1$	5.21	$\frac{0.6516}{0.6510} = 1$
$Rand$	1.5	1	2.1996	1	2.8384	1
EWD	1.5697	$\frac{0.4245}{0.2967} = 1.43$	2.3501	$\frac{0.4125}{0.2885} = 1.43$	3.0367	$\frac{0.3926}{0.2917} = 1.34$
$P_{fa} = 0.1,$	$N = M = 4$		$N = M = 6$		$N = M = 8$	
	$\sum_{i=1}^{N} G_i$	Υ	$\sum_{i=1}^{N} G_i$	Υ	$\sum_{i=1}^{N} G_i$	Υ
$rand - C$	2.9741	$\frac{0.8402}{0.6982} = 1.2$	4.2908	$\frac{0.8468}{0.5773} = 1.47$	5.1744	$\frac{0.8484}{0.5200} = 1.63$
$WSLR$ i.i.d	2.9949	$\frac{0.7348}{0.7329} = 1$	4.2882	$\frac{0.7222}{0.7218} = 1$	5.1679	$\frac{0.6464}{0.6441} = 1$
$WSLR$ Markov	3.03	$\frac{0.7597}{0.7595} = 1$	4.208	$\frac{0.7023}{0.7013} = 1$	5.107	$\frac{0.6388}{0.6385} = 1$
$Rand$	1.7640	1	2.5032	1	3.1552	1
EWD	1.8076	$\frac{0.4798}{0.3693} = 1.3$	2.6736	$\frac{0.4638}{0.3557} = 1.3$	3.4284	$\frac{0.4394}{0.3554} = 1.24$

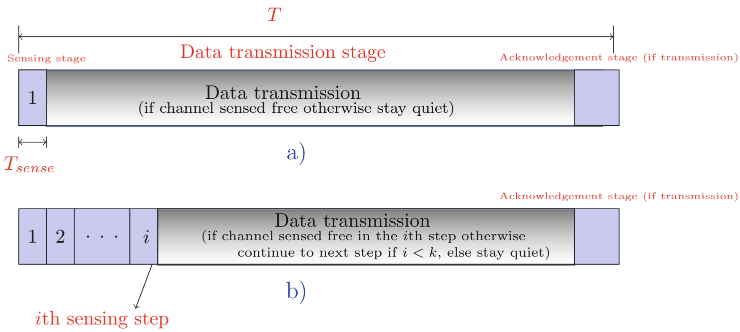


Fig. 1. Time slot structure: a) Single-channel sensing policy: When in a given time slot only one sensing step is available. b) Sequential-channel sensing policy: When in a given time slot more than one sensing step is available.

assumed to use a time-slotted system, and each primary user is either present in a channel for the entire time slot, or absent for the entire time slot [6,7].

The sensing stage in each time slot is divided into a number of sensing steps. Each sensing step is used by a CR to sense a different channel (see Fig. 1). In practice, improving the sensing accuracy implies increasing the sensing duration, whereby CRs may not be able to sense all the channels within the duration of a slot. We evaluate our proposed strategy for the scenario where the number of sensing steps k that a CR can utilize in a given time slot varies from 1 to M. If a CR finds a channel free in its ith sensing step, it transmits in that channel. However, if in all sensing steps channels are found to be busy, then the CR stays silent for the remaining duration of that time slot (see Fig. 1). When a free channel is found in the ith sensing step, the durations of the sensing stage and data transmission plus acknowledgement stage are iT_{sense} and $T - iT_{sense}$, respectively, where T_{sense} is the time required to sense each channel, T is the total duration of each slot and $T >> T_{sense}$. When multiple autonomous CRs search multiple potentially available channels, then from an individual CR perspective one of the following three events will happen in each sensing step: (1) It is the only one to find a channel free and transmit; the CR then has the channel for itself for the remainder of the time slot; (2) It finds that the channel is occupied by the PU or by another CR, then it continues looking in the next sensing step; (3) It visits a given channel, finds it free and transmits, but so does at least one other CR; a collision occurs. A CR infers that a collision has occurred whenever it fails to receive an acknowledgement (ACK) for a transmitted data frame. It should be noted that in this context, the channel sensing order $\mathbb{P}$, i.e., the order in which radios competing for the channels visit those channels, will affect their probability of successful access.

The vector of observation error probabilities is given by $\mathbf{e} = (P_{fa}, \sigma, \pi)$, where P_{fa} represents the probability of a false alarm, σ represents the probability that co-channel interference can be tolerated, i.e., when two CRs transmit simultaneously on the same channel, one of the transmissions is correctly decoded by its intended receiver with probability σ, and π represents the probability of a channel error.

3 An Adaptive WSLR Strategy

In this section, we propose an adaptive Win-shift, lose-randomize (WSLR) strategy, where adaptations are in the autonomous choice, by CRs, of the channel sensing order. Note that the sensing order that a CR employs can either come from the space of all permutations of M channels or from some subset thereof. However, the work in [14] shows that CRs can increase their average number of successful transmissions by adaptively selecting sensing orders from a predefined Latin Square of M channel indices. Note that when CRs select sensing orders from a Latin Square, $|\mathcal{S}| = M$, and two or more CRs can collide only if they select the same sensing order. We consider the case where each CR employs a common pre-defined sequence matrix (a Latin Square) Φ to select a sensing order in which k potential channels are to be visited in a given time slot, where k takes integer values between 1 to M.

For a given number of channels M there can be many Latin Squares. To select a sensing order from a common predefined Latin Square, CRs can employ any of the many Latin Squares. However, to make the analysis tractable, we assume that each CR employs a circulant matrix (which is an example of a Latin Square). For example, with $M = 3$, the circulant matrix Φ is given as:

$$\Phi = \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{pmatrix}$$

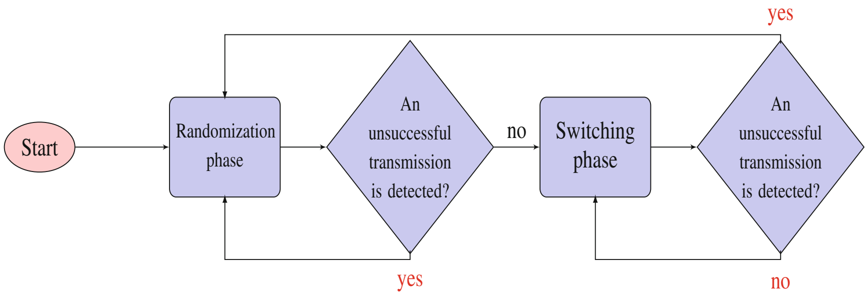


Fig. 2. The Win-shift, lose-randomize (WSLR) strategy.

The WSLR strategy is described in Fig. 2. The randomization and switching phases of the WSLR strategy (see Fig. 2) are explained as follows. In the *randomization phase*, each CR utilizes $\mathbf{p} = [\frac{1}{|\mathcal{S}|}, \frac{1}{|\mathcal{S}|}, \cdots, \frac{1}{|\mathcal{S}|}]$ ($| \mathcal{S} |$-element probability vector) for the selection of a sensing order. In the *switching phase*, the CR updates its current sensing order s_i to s_j where $j = (i \bmod | \mathcal{S} |) + 1$, i.e., in the next time slot, it shifts to the next sensing order to visit the channels.

The WSLR strategy is meant to address three aims:

1) Convergence: Note that when N CRs independently and randomly (with equal probability) select a sensing order (among $| \mathcal{S} | = N$ sensing orders) in each time slot, then the probability of arriving at orthogonal sensing orders in a time slot is $(1/N)^N \big(N!\big)$, and consequently the expected time required to arrive at orthogonal sensing orders is $N^N \big(1/N!\big)$. Clearly, this random strategy is inefficient as even when a CR attains a singleton status, i.e., the sensing order it has selected was not selected by another CR, it randomizes and with high probability it may lose the singleton status in the next time slot. In contrast to that, the WSLR strategy requires that singleton CRs should shift and non-singleton CRs should randomize. This reduces the number of CRs that randomly select a sensing order in the next time slot and hence increases the probability of arriving at orthogonal sensing orders. Hence with perfect private monitoring the WSLR strategy converges to conflict-free (orthogonal) sensing orders (provided that the number of CRs is less than or equal to M). For detailed proof of convergence of adaptation based sensing order selection strategies, we have results in [19, Theorem 4.1 and 4.2].

Unfortunately, when $\pi > 0$ (channel errors), in a given time slot if all the CRs have orthogonal sensing order selections, then an unsuccessful transmission (due to channel error) in any later time slot will lead a CR to erroneously move from shifting to randomization phase, in which case orthogonality may be lost, leading to reduced reward values. However, in Sect. 3.2 we show that for imperfect private monitoring our proposed strategy enables the CRs to increase their expected reward as compared to the random selection of sensing orders and other strategies.

2) Intertemporal sharing of rewards: Since different sensing orders may result in different rewards, intertemporal sharing of the sensing orders among autonomous CRs is achieved by allowing a CR to shift to the next sensing order if it has not observed an unsuccessful transmission, i.e., private outcome (U), in the previous time slot.

3) Discourage deviations: To discourage deviations, i.e., the CRs that select the sensing orders with higher rewards may prefer to again select those sensing orders in the next rounds, some punishment mechanism must be devised. This is achieved by triggering a switch to the randomization phase when an unsuccessful transmission is observed. Section 3.1 will further describe how the proposed mechanism discourages deviations by any of the autonomous CRs.

3.1 Analysis of the Adaptive WSLR Strategy

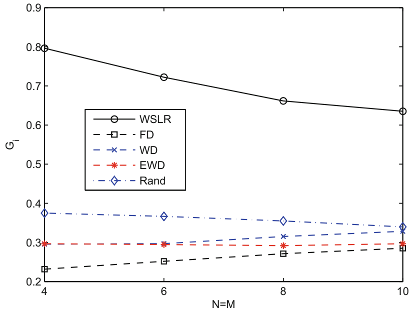
Without loss of generality, we assume that the channels are ordered by increasing probability of the PU being present, i.e., $\theta_1 \leq \theta_2 \leq, \cdots, \leq \theta_M$. For efficient channel utilization, we consider the scenarios where the N CRs utilize the N top rows of Φ for the selection of sensing orders. This is reasonable as the channel indices $1, 2, ..., M$ are ordered by increasing probability of the PU being present, hence the top N rows of Φ dominate in terms of having channels (in their initial columns) where PU's are less likely to be present. Note that for $N = M$, the entire matrix Φ is utilized by a CR for the selection of sensing orders. Let $\mathbf{S}_N$ represent the matrix of the top N rows of Φ.

Next, through extensive simulations, we analyze and compare the performance of the WSLR strategy for perfect and imperfect private monitoring.

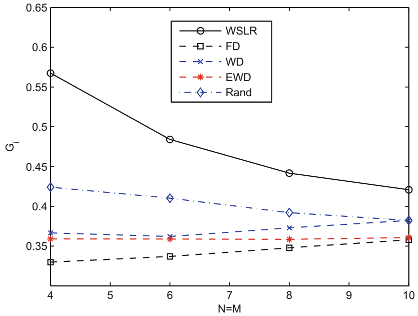
3.2 Simulation Results and Comparison with Other Strategies

Using simulation our aim is to compare the performance (e.g., in terms of total average reward per time slot in the CR network, expected reward of a CR per time slot, and the maximum envy ratio between a pair of CRs) of the WSLR strategy against: (1) When all CRs utilize random selection of sensing orders, *Rand strategy*; (2) the randomize after every collision (rand-C) strategy proposed in [6]. In the rand-C strategy, initially each CR independently and randomly (with equal probability) selects a sensing order. In the next time slots, a CR randomly (with equal probability) selects a new sensing order only if it has experienced a collision in the previous slot; otherwise, it retains the previously selected sensing order; and (3) An autonomous CR i considers deviating from the WSLR strategy while all other CRs follow the strategy. The studied deviations by the CR i are: (a) Always select the preferred sensing order $s_1 = (1, 2, ..., M)$, fixed deviation (FD); (b) Always select s_1 with probability $q = 0.75$ and s_2 with probability $(1 - q)$, weighted deviation (WD); and (c) Always select s_1 with probability $q = 0.75$ and $s_2, s_3, ..., s_N$ with probabilities $[\frac{(1-q)}{(N-1)}, \frac{(1-q)}{(N-1)}, .., \frac{(1-q)}{(N-1)}]$, extended weighted deviation (EWD). Moreover, we also evaluate the effect of varying the number of sensing steps on the performance of the proposed scheme. Note that calculations for expected reward per time slot of the CR i (G_i) are performed by Monte Carlo method using 15,000 Monte Carlo runs for dynamic channel selection process using different scenarios. $\Theta = (0.1, 0.2, 0.2, 0.3, 0.3, 0.5, 0.5, 0.5, 0.5, 0.5)$ is utilized PU duty cycle statistics vector (i.i.d PU occupancy model) for channels 1 to M respectively. For Markov Occupancy model, (0.6000, 0.6000, 0.6000, 0.6000, 0.6000, 0.5, 0.5, 0.5, 0.5, 0.5) is the vector of α_i's (see Sect. 2 for details) and (0.0660, 0.1500, 0.1500, 0.2600, 0.2600, 0.5, 0.5, 0.5, 0.5, 0.5) is the vector of β_i's for 1 to M channels respectively.

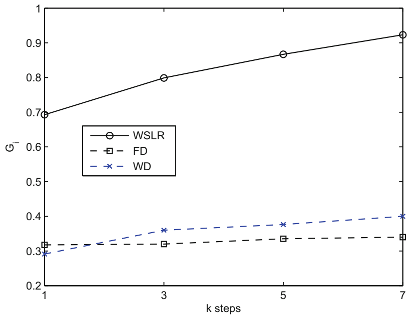
In Table 1, we evaluate different strategies in terms of the total average reward per time slot in the CR network and the highest envy ratio between a pair of CRs. Table 1 shows that the proposed method is not strongly affected by the stochastic model of the PU behavior. It also shows that the WSLR strategy performs equally well as the rand-C strategy in terms of maximizing the total average reward per time slot and performs significantly better in terms of ensuring envy-freeness among the competing CRs. It can also be seen from Table 1 that the random selection of sensing orders (rand) significantly reduces the total average reward per time slot in the CR network. When a CR considers deviating, while all other CRs follow the WSLR strategy, (see EWD in Table 1), it can be seen that there is no incentive in deviation from the WSLR strategy. Moreover, a deviating CR significantly reduces the total expected reward per time slot as compared to when all the CRs adopt the WSLR strategy. Figure 3a and b evaluate the expected reward per time slot achieved by a CR i using the different strategies under different scenarios, when CRs have perfect private monitoring and when their private monitoring is imperfect. From the two figures we can see that the WSLR strategy achieves the highest expected reward per time slot for the CR i as compared to when it considers deviating (while all the other CRs follow the WSLR strategy) and when all N CRs in the network utilize random selection of sensing orders. Note that in Fig. 3a and b, the loss in the expected reward of the CR i (when it adopts the WSLR strategy or random selection) is due to the non-homogeneity in channel availability statistics. The availability probabilities of the first 5 channels are at least 70 %, and the availability probabilities of the last five channels are 50 %. Hence, as $N = M$ increases, the expected reward of the CR i decreases, as with the increasing number of CRs the number of potentially available channels also increases but with high probability of a PU being present. Figure 3c evaluates the effect of varying the number of sensing steps on the performance of the different strategies in terms of expected reward of the CR i per time slot. It can be seen in Fig. 3c that when all the CRs with utilize the WSLR strategy then the expected reward per time slot of a CR increases as the number of sensing steps increase. However, it can also be seen in Fig. 3c that for the deviating CR (when all other CRs follow the WSLR strategy) there is little or no gain when more sensing steps are utilized for sensing. This is because for a given N and M, as k increases and when all the CRs utilize the WSLR strategy, it becomes more likely for a CR to find free channels in the later sensing steps as the WSLR strategy allows them to arrive at conflict-free allocations. When a CR deviates (while all other CRs stay on the WSLR strategy), the CR can either find



(a) Vector $\mathbf{e}$ is set to $\mathbf{e} = (P_{fa}, \sigma, \pi) = (0,0,0)$.



(b) Vector $\mathbf{e}$ is set to $\mathbf{e} = (0.1, 0.05, 0.05)$.



(c) Vector $\mathbf{e}$ is set to $\mathbf{e} = (P_{fa}, \sigma, \pi) = (0,0,0)$.

Fig. 3. a) and b) Expected reward per time slot of the CR i (G_i) as a function of $N = M$ CRs for different scenarios. c) Expected reward per time slot of the CR i as a function of the number of sensing steps k, with $N = 5$ CRs, $M = 9$ channels. PU occupancy model is i.i.d.

a free channel in its initial sensing steps (when the CR is the sole radio following this sensing order) as it selects with high probability the sensing order s_1 in which the availability probabilities of the first 2 channels are at least 90 %, or else it may collide with the other CRs (when some other CR also selects the same sensing order) and fail to find a free channel during that time slot. Hence, increasing the number of sensing steps has either little or no gain for the deviating CR i.

4 Concluding Remarks

We have studied the problem of coexistence and competition among multiple autonomous CRs for a shared pool of non-homogeneous spectrum resources. For efficient and fair co-existence, we present an adaptive WSLR strategy that does not require coordination from a centralized entity and utilizes noisy feedback (imperfect signals) to infer the actions of other CRs. To analyze the impact of imperfect signals of a CR on the performance of the proposed WSLR strategy, using simulations, we have explored the effects of false alarms, channel errors and co-channel interference tolerance on the performance of our proposed strategy. We have shown that the proposed strategy increases the average number of successful transmissions in the network and minimizes the payoff distribution conflict among competing CRs.

Acknowledgement. This work was funded by Academy of Finland under the Academy Post Doctoral funding grant of Dr. Zaheer Khan.

References

1. Schmidt, E.E., Mundie, C.: Realizing the full potential of government-held spectrum to spur economic growth (2012). http://www.whitehouse.gov/administration/eop/ostp/pcast
2. Bataller, H., Commission, E.: Promoting the shared use of radio spectrum resources in the internal market (2012). http://www.eesc.europa.eu
3. Melvin, J.: US regulators ok T-mobile testing of shared use of airwaves, 15 Aug 2012. http://www.reuters.com/article/2012/08/15
4. Ghasemi, A., Sousa, E.S.: Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs. IEEE Commun. Mag. **46**(4), 32–39 (2008)
5. Lipton, R. J., Markakis, E., Mossel, E., Saberi, A.: On approximately fair allocations of indivisible goods. In: Proceedings of the 5th ACM Conference on Electronic Commerce, EC 2004, pp. 125–131 (2004)
6. Anandkumar, A., Michael, N., Tang, A.: Opportunistic spectrum access with multiple users: learning under competition. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), pp. 1–9 (2010)
7. Liu, K., Zhao, Q.: Distributed learning in cognitive radio networks: multi-armed bandit with distributed multiple players. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 3010–3013 (2010)
8. Gai, Y., Krishnamachari, B., Jain, R.: Learning multiuser channel allocations in cognitive radio networks: a combinatorial multi-armed bandit formulation. In: Proceedings of the IEEE International Dynamic Spectrum Access Networks (DySPAN), pp. 1–9 (2010)

9. Khan, Z., Lehtomaki, J. J., DaSilva, L. A., Latva-aho, M., Juntti, M.: Adaptation in a channel access game with private monitoring. In: Proceedings of the Global Communications Conference, IEEE GLOBECOM 2013 (2013). Accepted for publication. http://www.ee.oulu.fi/~zaheer/Globecom2013.pdf
10. Khan, Z., Lehtomaki, J., DaSilva, L., Hossain, E., Latva-aho, M.: Opportunistic channel selection by cognitive wireless nodes under imperfect observations, limited memory: a repeated game model. IEEE Trans. Mob. Comput. **15**, 173 (2015)
11. Su, H., Zhang, X.: Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks. IEEE J. Sel. Areas Commun. **26**, 118–129 (2008)
12. Peh, E., Liang, Y.-C.: Optimization for cooperative sensing in cognitive radio networks. In: Proceedings of the IEEE International Wireless Communication Networking Conference (WCNC), pp. 27–32 (2007)
13. Liu, K., Zhao, Q., Krishnamachari, B.: Distributed learning under imperfect sensing in cognitive radio networks. In: Proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers (2010)
14. Khan, Z., Lehtomaki, J., DaSilva, L., Latva-aho, M.: Autonomous sensing order selection strategies exploiting channel access information. IEEE Trans. Mob. Comput. **12**(2), 274 (2013)