Opportunistic Channel Selection by Cognitive Wireless Nodes Under Imperfect Observations and Limited Memory: A Repeated Game Model

Zaheer Khan, Janne J. Lehtomäki, Luiz A. DaSilva, Ekram Hossain, *Fellow, IEEE*, and Matti Latva-Aho

Abstract—We study the problem of how autonomous cognitive nodes (CNs) can arrive at an efficient and fair opportunistic channel access policy in scenarios where channels may be non-homogeneous in terms of primary user (PU) occupancy. In our model, a CN that is able to adapt to the environment is limited in two ways. First, CNs have imperfect observations (such as due to sensing and channel errors) of their environment. Second, CNs have imperfect memory due to limitations in computational capabilities. For efficient opportunistic channel access, we propose a simple adaptive *win-shift lose-randomize (WSLR)* strategy that can be executed by a two-state machine (automaton). Using the framework of repeated games (with imperfect observations and limited memory), we show that the proposed strategy enables the CNs (without any explicit coordination) to reach an outcome that: 1) maximizes the total network payoff and also ensures fairness among the CNs; 2) reduces the likelihood of collisions among CNs; and 3) requires a small number of sensing steps (attempts) to find a channel free of PU activity. We compare the performance of the proposed autonomous strategy with a centralized strategy and also test it with real spectrum data collected at RWTH Aachen.

Index Terms—Autonomous cognitive nodes, adaptation, game theory, imperfect private monitoring, limited memory, non-homogeneous channels, and real spectrum occupancy data

1 Introduction

To help mitigate the critical stress on spectrum resources spurred by the ever more powerful smart devices, both a recent U.S. presidential advisory committee report and an FCC report recommend the use of spectrum sharing technologies [1], [2]. One potential solution is cognitive radio technology, in which cognitive nodes (CNs), such as small cell base stations (SCBSs), access points, and other types of nodes, are able to adapt intelligently to the environment through observation, exploration and learning. CNs utilize spectrum opportunistically by monitoring the licensed frequency spectrum to reliably detect primary user (PU) signals and operating whenever the PU is absent. The detection of PU signals can be achieved by: 1) spectrum sensing; 2) the use of geolocation databases; or 3) the combination of both [3].

In this paper, under the spectrum overlay model, we explore the question of how autonomous CNs with imperfect observations (such as due to sensing and channel errors and limited memory) can arrive at an efficient and fair opportunistic channel access policy in scenarios where multiple potentially available channels may offer different payoffs due to their non-homogeneity (in terms of PU occupancy). The real spectrum occupancy data collected at RWTH Aachen confirms that the spectrum resources are in general non-homogeneous in terms of PU occupancy. As an example, in Fig. 1 we illustrate average availability of a channel for a data set collected over the DECT bands by RWTH Aachen in different locations [4].

When selfish autonomous CNs compete for non-homogeneous potentially available channels to maximize their own payoffs, then a channel access strategy that is not carefully designed can degrade individual and total network throughput. Let us consider an example scenario where N cognitive small cell base stations, which we can refer to as cognitive nodes, are deployed by multiple independent wireless operators in the same area. These are within the interference range of one another and compete for a channel out of M potentially available channels. Since N CNs belong to different service providers they require an opportunistic spectrum access (OSA) strategy that does not rely on a common coordinator. The CNs are required to perform periodic spectrum sensing so that when a PU becomes active in a channel they can vacate that channel. Hence, if one exclusive channel is considered equivalent to one unit of throughput then the maximum throughput achieved by a CN that selects channel i can be no more than $(1-\theta_i)$, where θ_i is the duty cycle of the PU that operates in channel i. Without loss of generality, let us order the channels according to their duty cycles, such that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_M$. When two or more CNs select their channel autonomously, then a selfish CN would want to select channels with low PU duty cycles in an effort to obtain a higher throughput. Obviously, two or more CNs that implement such always "select

- *Z. Khan, J. Lehtomäki, and Matti Latva-Aho are with the Centre for Wireless Communications, University of Oulu, Finland. E-mail: {zaheer, jannel, matla}@ee.oulu.fi.*
- *L. A. DaSilva is with the Telecommunications Research Centre, Trinity College Dublin, Ireland, and Virginia Tech, Arlington, VA 24061. E-mail: ldasilva@vt.edu.*
- *E. Hossain is with the Department of Electrical and Computer Engineering, University of Manitoba, Canada. E-mail: Ekram.Hossain@umanitoba.ca.*

Manuscript received 27 May 2014; revised 9 Feb. 2015; accepted 5 Mar. 2015. Date of publication 13 Mar. 2015; date of current version 1 Dec. 2015.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2015.2412940

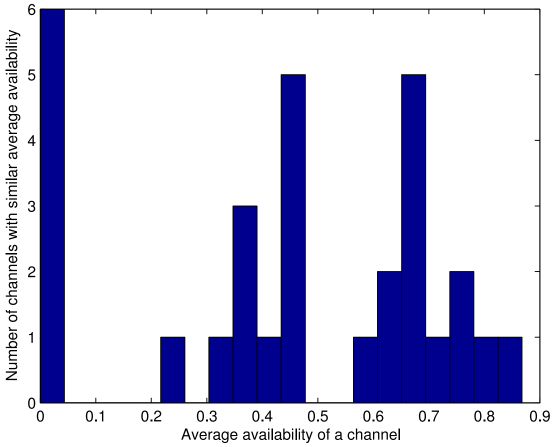


Fig. 1. Average availability of a channel for a data set collected over the DECT bands in different locations by RWTH Aachen [4].

channels with low PU duty cycle" strategy may simultaneously sense free the same channels more often and decide to transmit on the same channels. Consequently, collisions will occur and many CNs will have an unsuccessful channel access.

Autonomous competition for potentially available channels involves interaction between autonomous self-motivated CNs. The framework of repeated games provides useful tools to analyze conflicts among players that interact repeatedly over time [5], [6]. Autonomous CNs operating in a CN network are unsure about when precisely their interactions will end. Therefore, the model of repeated games with an infinite time horizon can be used to analyze such situations. In our work, we utilize the framework of the repeated games with imperfect private monitoring and limited memory for the study of channel sensing/access order selection among autonomous CN nodes [6]. In our model of channel sensing/access order selection, each CN has following limitations: 1) Each CN can only utilize their own feedback information (such as occurrence of collisions) to find a way to autonomously arrive at those channel sensing/access orders that minimize the likelihood of collisions among them. 2) Each CN has false alarms in their sensing observations of primary usage of channels and also their own feedback is noisy. For instance, due to channel errors, feedback (such as occurrence of collision information) from their corresponding receiving nodes can be imperfect. 3) They have imperfect memory regarding outcomes of their past channel order selection decisions due to limitations in storage and computational capabilities. In particular, they use only simple adaptive strategies based on the outcomes during the previous period of play.

The main contributions of this paper can be summarized as follows:

- We propose and evaluate the performance of a simple adaptive win-shift lose-randomize (WSLR) strategy for opportunistic channel selection among autonomous CNs and test it with real spectrum occupancy data collected at RWTH Aachen [4].
- To study the performance of the proposed strategy against selfish deviations, we formulate the opportunistic spectrum access problem as a repeated game with imperfect observations and limited memory. We show that our proposed strategy can be executed by only a two-state machine (automaton). We also show that all CNs playing the WSLR strategy constitutes an equilibrium.
- Using analytical and simulation results, we compare the performance of our proposed strategy against other existing strategies. The proposed strategy enables the CNs to maximize the total average payoff in the network and also ensures fairness by leading the autonomous CNs to engage in inter-temporal sharing of the rewards from cooperation. It reduces the likelihood of collisions among CNs and hence reduces the costs due to retransmission attempts. Moreover, it requires a small number of sensing steps (attempts) to find a channel free of PU activity.
- We compare the performance of the proposed strategy against a centralized allocation of potentially available channels. With chosen system parameters, the performance of the proposed scheme is observed to be near the optimum performance obtained by a centralized allocation.
- We consider both single and sequential channel sensing policies to evaluate the effect of varying the number of channels that a CN can sense on the performance of our proposed adaptive strategy.[1]

The rest of this paper is organized as follows. Section 2 reviews the relevant literature addressing the problem of autonomous channel selection for OSA. The system model and assumptions are stated in Section 3. Section 4 introduces the proposed repeated game model for opportunistic channel selection. The WSLR strategy and its complexity are presented in Section 5. In Section 6, we present the analytical results for the proposed scheme. In Section 7, we present the simulation results and also compare our proposed WSLR strategy to related strategies proposed in other work in the literature. Finally, Section 8 concludes the paper. A list of important symbols used in the paper is shown in Table 1.

2 Related Work and Motivation

In this section, we explain the motivations behind our work and also summarize the main differences between the related works and our own.

2.1 Opportunistic Channel Selection with Cooperative Protocol Followers

In sensing-based OSA, when CNs compete for multiple potentially available channels, time-slotted multiple access has generated much interest [7], [8]. In CN networks, to protect a PU from harmful interference, the CNs are required to perform periodic spectrum sensing so that when a primary user becomes active in a channel, the CNs can vacate that channel [7], [9], [10], [11]. Single-channel

1. Under a single-channel sensing policy, in any given time slot a CN first selects a channel to sense and transmits if that channel is free; otherwise, it stays silent for the entire duration of that time slot. Under a sequential-channel sensing policy, a CN can sense more than one channel within the duration of a time slot. These policies will be discussed in detail in Section 3.1.

TABLE 1
List of (Important) Symbols Used in This Paper and Their Meaning

Symbol	Meaning	Symbol	Meaning
N	Number of autonomous CNs	$\mathcal{N}$	Set of autonomous CNs
M	Number of potentially available channels	$\mathcal{M}$	Set of potentially available channels
θ_i	Probability of PU being present in channel i	T_L	Total duration of time slot
T_{sense}	Time required to sense a channel	P_d	Detection probability
$\mathcal{S}$	Set of sensing orders	s_i	A channel sensing order i
Φ	Sequence Matrix	k	Number of sensing steps
Ξ	Set of possible outcomes (signals)	ξ_i	Private outcome (signal) of CN i
$g_i\big((s_i, \mathbf{s}_{-i})\big)$	CN i's average reward in the stage game (perfect monitoring)	$g_i\big((s_i, \mathbf{s}_{-i}), \mathbf{e}\big)$	CN i's average reward in the stage game (imperfect monitoring)
$G_i(\mathbf{S}_a, \mathbf{S}_\xi)$	CN i's average reward per time slot in the repeated game	γ_g	gain parameter
ε_{ij}	*Envy-ratio* of CN i for CN j	Υ	*Highest average envy ratio* between any pair of CNs
$P^*_{s,1}$	Probability of finding a channel free in first step (given that the channel is free)	E[TTO]	Average number of time slots required to orthogonal sensing orders

sensing policies are investigated in [7], [10], [11] while sequential-channel sensing policies are described in [12], [13], [14], [15].

Jiang et al. in [13] investigate the optimal selection of a channel sensing order for a single cognitive node. In contrast to that, our work considers competition for channels among multiple CNs. Distributed learning and allocation strategies for multichannel cognitive medium access are considered in [7]. The problem of learning multiuser channel allocation for cognitive medium access is formulated as a combinatorial multiarmed bandit problem in [11]. The work in [11] considers the scenario where due to geographic dispersion, each secondary user can potentially see different primary user occupancy behavior on each channel. A channel allocation solution is proposed in [11] that maximizes the expected sum throughput. The work in [10] considers the problem of distributed learning and channel allocation in cognitive radio networks under imperfect sensing. The objective of the works in [7], [10], [11] is to maximize the total system throughput using a fair decentralized policy, and different from our work, they assume that each CN cooperatively follows the same strategy (protocol).

2.2 Opportunistic Channel Selection With Self-Motivated Users

When channel rewards are non-homogeneous, it is possible that some CNs may deviate from a cooperative protocol and behave selfishly to maximize their own usage at the expense of the total system throughput. To analyze the conflict among multiple self-motivated users, many game theoretic research papers have been published related to channel selection. However, most of the works assume that players can observe the other players' payoffs or actions [16], [17]. Moreover, typical models of repeated game for studying competition in OSA also assume that players have unlimited memory. In such repeated game models, players condition their strategies on the entire history of the game, irrespective of how long and complicated that history may be. Yet, in reality, such assumptions are unlikely to hold in autonomous CN networks. Autonomous CNs operating in a network are not able to observe the actions of all other CNs. They need to infer the actions of other CNs based on noisy feedback from the wireless environment. Moreover, due to computational limitations, autonomous CNs may not have the ability to remember and take into account in their decisions the full history of a game.

To address such challenges, we utilize the framework of repeated games with imperfect private monitoring for the study of channel sensing/access order selection among autonomous CNs. Note that a repeated game model with imperfect private monitoring is different from a game with incomplete information and as well as from a game with imperfect information. Players in an incomplete information game might not have common knowledge of payoffs, who the other players are, and what moves are possible. Players in a game of imperfect information are simply unaware of the actions chosen by other players. However, they know the preferences/payoffs of these other players. In an imperfect private monitoring game, players do not have information regarding payoffs of other players and also they do not have perfect observation of other players' actions. In our imperfect private monitoring model for dynamic channel selection, each autonomous CN is required to infer the actions of other CNs based on noisy feedback (such as occurrence of collision information) from their corresponding receiving nodes.

The three works in [18], [19], [20] utilize the framework of repeated games with imperfect public monitoring in the context of channel selection. However, in our work we consider the framework of repeated games with imperfect *private* monitoring. In repeated games with imperfect public monitoring, players cannot observe the other players' actions directly, but can observe imperfect public signals about them, i.e., they have common knowledge about such actions. In repeated games with imperfect private monitoring, players cannot observe the other players' actions directly, but can only observe private signals about them, i.e., they do not have common knowledge about such actions. The three other main differences between these works and our work are: 1) The works in [18], [19], [20] consider a single resource (channel), while our work is related to sensing order/channel access selection for the scenarios where multiple channels are potentially available for spectrum access. 2) The proposed method in [18], [19], [20] involves a monitoring unit measuring interference/congestion level in the channel, which is then sent as a binary public signal to the players or it involves some information exchange among the players. Our proposed method does not require such information exchange. The only information our method requires is that initially, each CN obtains the statistics of primary users' duty cycles and the number of other active CNs through a spectrum access system (SAS) [1], [21]. Later there is no information exchange. 3) In [18], [19], [20], a player's memory at period t is a collection of public feedback signals and its own actions from period 0 to $(t-1)$. In our work, the proposed strategy requires only one-round of memory, as it responds to the previous round of play. Only a few works in the literature utilize the framework of repeated games with imperfect private monitoring to model wireless communications and networking problems. The works in [22], [23], [24] study packet forwarding in multi-user wireless networks as a game with imperfect private monitoring. In the rest of the paper we use the terms "imperfect observations" and "imperfect monitoring" interchangeably.

In [25], the authors propose an indirect reciprocity game model for opportunistic channel access. However, different from our work, in [25] CNs help primary users by choosing a power level to relay their (primary users') information. Based on this the CNs gain reputations which in turn determine how much they can access a certain amount of vacant licensed bandwidth in the future. Moreover, unlike our work, in [25] time is divided into two slots, where the PU transmits in the first slot and a base station updates reputation information of the CNs in the second slot. The work in [26] proposes a new game called the Chinese restaurant game and in [27] the authors study its application to cognitive radio networks. In [27] the authors use the Chinese restaurant game model to study the problem of learning and access in cognitive radio networks. However, different from our work, the work in [27] considers a scenario where there is a log file in the server of the secondary network which records each CN's channel belief and channel selection results. The authors in [28] study the problem of learning and access in opportunistic spectrum access as a Chinese restaurant game. However, different from our work, the authors consider the scenario where a time slot is further sub-divided into three sub-slots: N CNs simultaneously perform sensing in the first sub-slot. In the second sub-slot, CNs sequentially make their channel access decisions based on the information they collected and report their decisions as well as their sensing results via a dedicated common control channel which can be overheard by all other CNs. Finally, CNs transmit their data through the channels they selected in the third sub-slot. Note that in our work we focus on autonomous scenarios where there is no information exchange among CNs.

In this paper, we build on our previous research on autonomous channel selection (with perfect observations) for CNs [29]. Unlike the work in [29], we assume that the CNs have imperfect observations. Also, unlike the work in [29], in this paper, to take into account the impact of PU occupancy behavior on the performance of the proposed adaptive channel selection scheme, and we evaluate the performance of the proposed scheme by testing it with real spectrum occupancy data collected at RWTH Aachen [4].

3 System Model and Assumptions

3.1 Network Model

We assume a multichannel CN network in which N autonomous CNs have M potentially available channels. Let $\mathcal{M} = \{1, 2, \ldots, M\}$ represent the set of (potentially available) channels and $\mathcal{N} = \{1, 2, \ldots, N\}$ represent the set of autonomous CNs. To make the analytical results tractable, we utilize the i.i.d. model of PU channel occupancy for theoretical analysis. In the i.i.d. model of PU channel occupancy (also adopted in [7]), 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. In Section 7, we test the proposed strategy with both real spectrum data collected at RWTH Aachen and the i.i.d. model of PU channel occupancy and find that the performance of the proposed adaptive strategy is not strongly affected by PU behavior.

Without loss of generality, we order the channels by the increasing probability of the PU being present, i.e., $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_M$. We assume that the CNs are unable to distinguish between a PU and other CN transmissions. In our work, the statistics of primary user duty cycle are assumed to be known to the autonomous CNs. In practice, the autonomous CNs may obtain the statistics of primary user duty cycle through the use of geo-location databases [1]. For instance, to protect the PUs from interference and to aid users seeking to utilize the spectrum for secondary usage, recent approaches to spectrum sharing have suggested the use of a spectrum manager entity, such as a Spectrum Access System [1], [21]. In SAS-based systems, multiple independent users are required to register with the SAS before becoming active in the network. The SAS system may maintain a geo-location database that contains records of the statistics of the primary users' duty cycles. When a CN registers with the SAS, the latter can communicate this information to the CN. Each CN can sense only one channel at a time, and due to hardware constraints, at any given time each CN can either sense or transmit, but not both.

3.2 Single and Sequential Channel Sensing Policies

The primary users and CNs are both 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 [7], [30]. The CNs use the beginning of each slot to sense the channels in some order (based on their sensing order selection strategies, as will be explained in Sections 4 and 5) to find a channel that is free of PU (or other CN) activity. We refer to this as the sensing stage (see Fig. 2). The CN then accesses the first vacant channel it finds, if one exists. We refer to this as the data transmission stage. Let $\mathcal{S}$ denote the set of sensing orders. Note that the sensing order that a CN employs can either come from the space of all permutations of M channels, or from some subset thereof.

In this paper, we consider both single and sequential channel sensing policies to evaluate the effect of varying the number of channels that a CN can sense on the performance of our proposed adaptive strategy. In other words, we

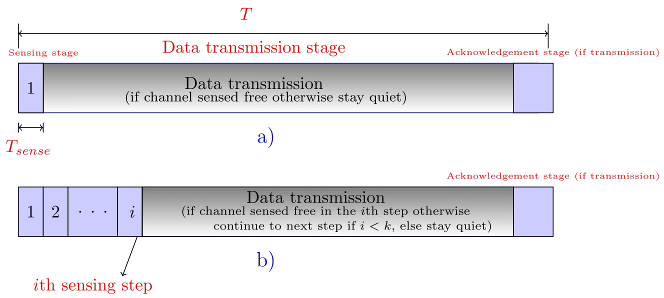


Fig. 2. Time slot structure with sensing, data transmission and acknowledgement stages. a) Single-channel sensing policy: when in a given time slot only one sensing step is available. b) Sequential-channel sensing policy: when in a given time slot more than one sensing step is available.

evaluate our proposed strategy for the scenario where the number of sensing steps k that a CN can utilize in a given time slot varies from 1 to M. When the CN can sense more than one channel in a time slot then the sensing stage in each slot is divided into a number of sensing steps. Each sensing step is used by a CN to sense a different channel. If a CN 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 CN stays silent for the remaining duration of that time slot (see Fig. 2). In practice $T_L \gg T_{sense}$, where T_{sense} is the time required to sense each channel, and T_L is the total duration of each slot.

In our work, we consider the conventional energy detector for channel sensing. To protect the transmissions by the incumbent, the detection probability $(P_{d,i})$ of an autonomous CN 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 [31]. In the literature this is defined as the constant detection rate (CDR) requirement. For a fixed target detection probability, the probability of false alarm of a CN is a variable. Current sensing mechanisms proposed for opportunistic access in [32], [33] have been shown to achieve the target detection probability. Moreover, the works in [32], [33] also show how the probability of false alarm of each CN i for the targeted $P_{d,i} = P_d$ can be approximated.

3.3 Sensing Order Sequence Matrix

In our model, each autonomous CN independently selects a channel sensing order from a pre-defined sequence matrix Φ in which k potential channels are to be visited in a given time slot, where k takes integer values between 1 to M. The pre-defined sequence matrix Φ that a CN employs can either come from the space of all permutations of M channels, or from some subset thereof such as a Latin Square. A Latin Square is an M by M matrix whose entries consist of M symbols such that each symbol appears exactly once in each row and each column. Note that when CNs select sensing orders from a Latin Square, $|\mathcal{S}| = M$, and two or more CNs can collide only if they select the same sensing order. In our work, each autonomous CN employs the same sequence matrix Φ which is a Latin Square of M channel indices as we proposed in [15]. However, using the proposed WSLR strategy, the way these channel sensing orders are adaptively selected (from the employed sequence matrix Φ) by an autonomous CN is different from the method proposed in [15]. The primary reason for employing the same sequence matrix as in [15] is as follows: We showed in [15] that the adaptive selection of sensing orders from a predetermined subset of permutations of M channels (such as a Latin Square of M channel indices) leads to faster convergence to those sensing orders that minimize the likelihood of collisions among the CNs, as compared to when the sensing orders are selected from the large space of all permutations of M channels. This in turn increases the average number of successful transmissions of each autonomous CN.

3.4 Imperfect Observations and Channel Sensing/Access Order Selection

In our model, to autonomously arrive at conflict-free sensing orders each CN needs to infer the actions of other CNs (whether the other CNs have selected the same sensing order) based on noisy feedback from the wireless environment. Conflict-free sensing orders are those in which two or more CNs never simultaneously sense the same channels and therefore never collide with one another. A CN infers that in a given time slot it has selected a conflict-free sensing order if it does not experience an unsuccessful communication when it adopts the sensing order. When it experiences an unsuccessful communication, i.e., it fails to receive an ACK for a transmitted data frame, it infers that there is a conflict with some other CR. However, in practice, feedback observations in a wireless environment could be imperfect. Next we explain the impact of false alarm, channel error, and the capture effect on the channel sensing/access order selection decision of a CN.

A false alarm would have the effect of a CN thinking a channel is busy when it is in fact free of both PU and other CN activity. Moreover, a false alarm could also have the effect of a CN thinking that the two CNs have selected conflict-free sensing orders when in fact they may have selected the same sensing order (see Fig. 3, scenario b). This may slow down the convergence to conflict-free sensing orders.

A channel error would lead to failure in receiving an ACK for a transmitted data frame. This can have the effect of a CN thinking that the two CNs have selected the same sensing order when in fact they may have selected conflict-free sensing orders (see Fig. 3, scenario d).

In wireless communications, the capture effect, also called co-channel interference tolerance, is the ability of a radio receiver to receive a signal from one transmitter in the presence of interference from one or more other transmitters [34]. When more than one CN decides to transmit on the channel, the data frame reception may still succeed, i.e., an ACK is received, due to the co-channel interference tolerance. Co-channel interference tolerance would have the effect of a CN thinking that the two CNs have selected conflict-free sensing orders when in fact they have selected the same sensing order (see Fig. 3, scenario c). This may slow down the convergence to conflict-free sensing orders.

4 Opportunistic Channel Sensing/Access Order Selection Game

4.1 Repeated Game Model

We now formally define an opportunistic channel sensing/access order game with imperfect observations. CN i, where

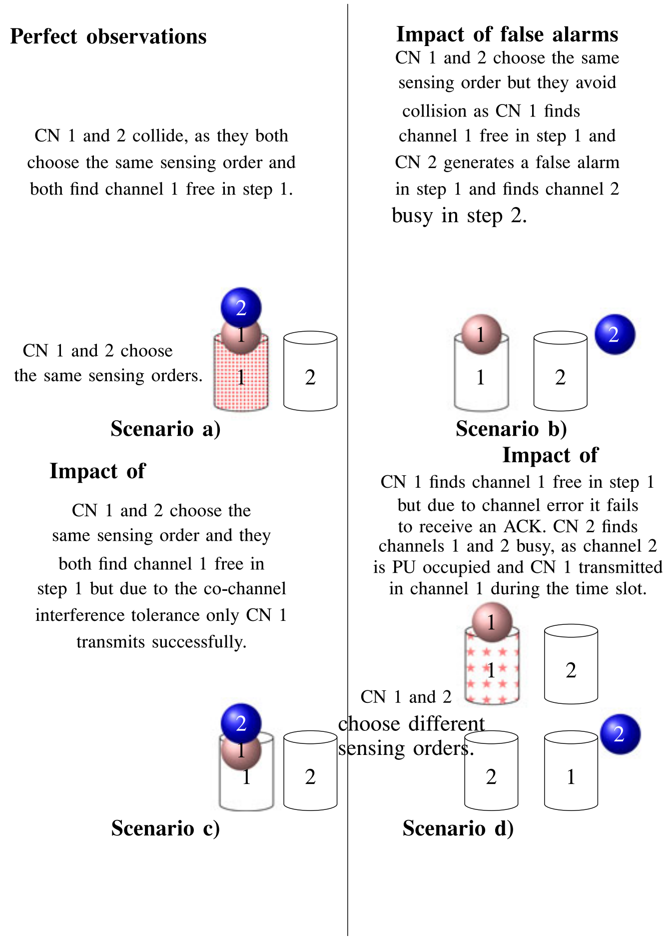


Fig. 3. Example scenarios illustrating the impact of imperfect observations on opportunistic multichannel access using pre-selected sensing orders. The cylinders represent the sensing orders, a cylinder with dots represents a collision between CN 1 and 2, and a cylinder with stars represents failure to receive an ACK for a transmitted data frame due to impairments to the wireless channel.

$i \in \mathcal{N}$, repeatedly plays the channel selection game over an infinite time horizon, $t = 0, 1, \ldots$. In each stage (corresponding to a time slot), CN i chooses a sensing order $s_i \in \mathcal{S}$ to sense the channels sequentially for spectrum opportunities, where $\mathcal{S}$ is the set of sensing orders.

In a given time slot, CNs searching for spectrum opportunities face one of the following outcomes: successful transmission (transmits and receives ACK), unsuccessful transmission (transmits but receives no ACK), or no transmission (when all channels sensed by that CN were found busy). At the end of each stage, a CN observes a private signal $\xi_i \in \Xi$. The action s_i and ξ_i are CN i's private information.

A CN i's expected payoff in the stage game is given by

$$g_i\big((s_i, \mathbf{s}_{-i})\big) = \sum_{\xi_i \in \Xi} u_i(s_i, \xi_i) p(\xi_i \,|\, (s_i, \mathbf{s}_{-i})), \qquad (1)$$

where s_i is the action of CN i, $\mathbf{s}_{-i}$ is the action profile of all other CNs, and $u_i(s_i, \xi_i)$ is the realized reward of CN i. $u_i(s_i, \xi_i)$ equals 1 if using sensing order s_i CN i transmits a frame and receives an ACK for that frame, i.e., $\xi_i = T$, otherwise it is 0, and $p(\xi_i \,|\, (s_i, \mathbf{s}_{-i}))$ is the conditional probability of signal ξ_i, where $\xi_i \in \Xi$, with $\Xi = \{$Unsuccessful transmission (U), Successful transmission (T), Channels found busy $(B)\}$.

For a CN i, in a repeated game, there is an infinite sequence of joint actions $(s_i^1, \mathbf{s}_{-i}^1), (s_i^2, \mathbf{s}_{-i}^2), \ldots$ and observed signals $\xi_i^1, \xi_i^2, \ldots$. The average reward per round of CN i over the infinite sequence of joint actions is given by

$$G_i(\mathbf{S}_a, \mathbf{S}_\xi) = \lim_{\bar{T} \to \infty} \frac{1}{\bar{T}} \sum_{t=1}^{\bar{T}} u_i\big((s_i^t, \mathbf{s}_{-i}^t), \xi_i^t\big), \qquad (2)$$

where $u_i\big((s_i^t, \mathbf{s}_{-i}^t), \xi_i^t\big)$ is the reward of CN i when joint action $(s_i^t, \mathbf{s}_{-i}^t)$ is played and signal ξ_i^t is observed, $\mathbf{S}_a = (s_i^t, \mathbf{s}_{-i}^t)_{t=0}^{\bar{T}}$, and $\mathbf{S}_\xi = \big(\xi_i^t\big)_{t=0}^{\bar{T}}$ are the sequences of action profiles and signals respectively.

Let $\mathbf{e} = (P_{fa}, \sigma, \pi)$ represent the vector of observation error probabilities, where P_{fa} represents the probability of a false alarm, π represents the probability of a channel error, i.e, when a CN transmits, its transmission is not correctly decoded by its intended receiver with probability π, and σ represents the probability that co-channel interference can be tolerated, i.e., when two or more CNs transmit simultaneously on the same channel, only one of the transmissions is correctly decoded by its intended receiver with probability σ. When observation error probabilities are taken into account, the expected payoff of CN i in the stage game is $g_i\big((s_i, \mathbf{s}_{-i}), \mathbf{e}\big)$.

4.2 Illustrative Examples

We show two examples where $N = 2$ autonomous CNs operate in $M = 2$ potentially available channels. In Sections 6 and 7 we provide detailed results for the scenarios where $N, M > 2$.

4.2.1 *Two CNs with Perfect Observations* $(\sigma = \pi = P_{fa} = 0)$

Consider the case in which $N = 2$ autonomous CNs have $M = 2$ potentially available channels. The reward table for the two-CN, two-channel selection stage game with (perfect observations), where $s_1 = (1, 2)$ and $s_2 = (2, 1)$ represent the two channel sensing orders, is given as:

	CN 2 plays s_1	CN 2 plays s_2
CN 1 plays s_1	0, 0	$(1-\theta_1), (1-\theta_2)$
CN 1 plays s_2	$(1-\theta_2), (1-\theta_1)$	0, 0

This case reduces to the well-known battle of the sexes game [35], and it is simple to prove that the game admits two pure strategy and one mixed strategy Nash equilibria. The vectors (s_1, s_2) and (s_2, s_1) are both pure strategy equilibria but, for $\theta_1 < \theta_2$, CN 1 prefers the first and CN 2 prefers the second (see reward table). The mixed strategy equilibrium is given by the *equalizing strategies* $\mathbf{p_1} = \big(\frac{(1-\theta_1)}{(2-\theta_1-\theta_2)}, \frac{(1-\theta_2)}{(2-\theta_1-\theta_2)}\big)$ and $\mathbf{p_2} = \big(\frac{(1-\theta_2)}{(2-\theta_1-\theta_2)}, \frac{(1-\theta_1)}{(2-\theta_1-\theta_2)}\big)$, where $\mathbf{p_1}$ and $\mathbf{p_2}$ are probability mass functions assigned by CNs 1 and 2 over their action spaces $\mathcal{S}$. The equalizing strategy is a strategy that produces the same average payoff no matter what the opponent does.

We now explore the impact of false alarm, channel error, and co-channel interference tolerance on the expected payoff of the two-CN sensing order selection game.

4.2.2 Two CNs with Imperfect Observations

We investigate the impact of imperfect observations on the expected reward of an autonomous CN under different scenarios such as i) CNs with false alarms but no channel errors and no co-channel interference tolerance. The expected payoff for this scenario is given as in (3) below

$$g_i\big((s_i, s_{-i}), \mathbf{e}\big) = \begin{cases} (1-\theta_1)(1-P_{fa})P_{fa} + (1-\theta_1)(1-P_{fa})P_{fa}(1-\theta_2)(1-P_{fa}) \\ +(1-\theta_1)P_{fa}P_{fa}(1-\theta_2)(1-P_{fa})P_{fa} + \theta_1(1-\theta_2)(1-P_{fa})P_{fa}, \\ \quad \text{if } (s_i, s_{-i}) = (s_1, s_1) \text{ or } (s_i, s_{-i}) = (s_2, s_2) \\ ((1-\theta_1) + (1-\theta_1)(1-\theta_2)P_{fa}^2 + \theta_1(1-\theta_2)P_{fa})(1-P_{fa}), \\ \quad \text{if } (s_i, s_{-i}) = (s_1, s_2) \\ ((1-\theta_2) + (1-\theta_2)(1-\theta_1)P_{fa}^2 + \theta_2(1-\theta_1)P_{fa})(1-P_{fa}), \\ \quad \text{if } (s_i, s_{-i}) = (s_2, s_1); \end{cases} \tag{3}$$

and ii) CNs with false alarms, channel errors and co-channel interference tolerance. The expected payoff for this scenario is given as in (4),

$$g_i\big((s_i, s_{-i}), \mathbf{e}\big) = \begin{cases} ((1-\theta_1)P_{fa}(1-\pi) + (1-\theta_1)P_{fa}(1-\theta_2)(1-\pi) \\ +\theta_1(1-\theta_2)P_{fa}(1-\pi) + (1-\theta_1)(1-P_{fa})\sigma/2 \\ +\theta_1(1-\theta_2)(1-P_{fa})\sigma/2 + (1-\theta_1)P_{fa}^3(1-\theta_2)(1-P_{fa})(1-\pi) \\ +(1-\theta_1)P_{fa}^2(1-\theta_2)(1-P_{fa})\sigma/2)(1-P_{fa}), \\ \quad \text{if } (s_i, s_{-i}) = (s_1, s_1) \text{ or } (s_i, s_{-i}) = (s_2, s_2) \\ (1-\theta_1)(1-P_{fa})(1-\pi) + (1-\theta_1)P_{fa}(1-\theta_2)(1-P_{fa})P_{fa}(1-\pi) \\ +\theta_1(1-\theta_2)(1-P_{fa})P_{fa}(1-\pi), \\ \quad \text{if } (s_i, s_{-i}) = (s_1, s_2) \\ (1-\theta_2)(1-P_{fa})(1-\pi) + (1-\theta_2)P_{fa}(1-\theta_1)(1-P_{fa})P_{fa}(1-\pi) \\ +\theta_2(1-\theta_1)(1-P_{fa})P_{fa}(1-\pi), \\ \quad \text{if } (s_i, s_{-i}) = (s_2, s_1). \end{cases} \tag{4}$$

For example in Eq. (4), when the CN i selects sensing order s_1 and the CN $-i$ selects s_2 then the user i can observe the signal T and get $u_i(s_1, T) = 1$ (see Eq. (1)) when: 1) In step 1, the channel 1 is free from the primary user, the user does not generate a false alarm in sensing, the user transmits and its transmission is correctly decoded by its intended receiver. The probability of this is given by the first term of Eq. (4), when $(s_i, s_{-i}) = (s_1, s_2)$. 2) In step 1, the channel 1 is free from the primary user, but the user generates a false alarm in sensing. In step 2, the channel 2 is free from the primary user, user $-i$ generates a false alarm in channel sensing in its first step, the user i does not generate a false alarm in its sensing step 2, the user transmits and its transmission is correctly decoded by its intended receiver. The probability of this is given by the second term of Eq. (4), when $(s_i, s_{-i}) = (s_1, s_2)$. 3) In step 1, channel 1 is occupied by the primary user. In step 2, channel 2 is free from the primary user, user $-i$ generates a false alarm in sensing in its first step, user i does not generate a false alarm in its sensing step 2, the user transmits and its transmission is correctly decoded by its intended receiver. The probability of this is given in the third term of Eq. (4), when $(s_i, s_{-i}) = (s_1, s_2)$.

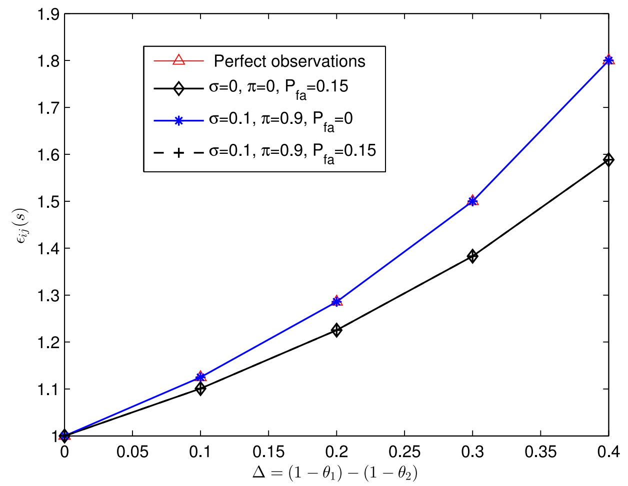


Fig. 4. Envy ratio as a function of Δ for different scenarios. The probability of PU being present (in preferred channel) is set to $\theta_1 = 0.1$ and θ_2 is varied.

In Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2015.2412940, available online, we illustrate an example of the probability tree diagram method used in deriving Eqs. (3) and (4). Note that Eq. (4) reduces to Eq. (3) when $\sigma = \pi = 0$. Moreover, σ, π and P_{fa} are assumed to be unknown to the CNs.

4.3 Envy-Ratio in the Proposed Game

We study the problem of efficient and fair utilization of potentially available channels that may offer different payoffs due to their non-homogeneity. The concept of fairness that we focus on is envy-freeness [36]. An outcome is envy-free if no CN prefers the expected payoff of another CN to its own, i.e., an envy-free outcome equalizes everyone's payoffs.

We next define the *envy-ratio* of CN i for CN j as follows.

Definition 1. *In an action profile* $\mathbf{s}$, *the* envy ratio *of CN i for CN j is the ratio of the reward obtained by j to the reward obtained by i. It is given as*

$$\varepsilon_{ij}(\mathbf{s}) = \frac{g_j\big((s_j, \mathbf{s}_{-j}), \mathbf{e}\big)}{g_i\big((s_i, \mathbf{s}_{-i}), \mathbf{e}\big)}, \quad \text{for } g_i\big((s_i, \mathbf{s}_{-i}), \mathbf{e}\big) > 0. \tag{5}$$

In the repeated game, the *average envy ratio* of CN i for CN j is given by

$$\Upsilon_{ij}(\mathbf{S}_a, \mathbf{S}_\xi) = \frac{G_j(\mathbf{S}_a, \mathbf{S}_\xi)}{G_i(\mathbf{S}_a, \mathbf{S}_\xi)}. \tag{6}$$

The *highest average envy ratio* between any pair of CNs is given as

$$\Upsilon(\mathbf{S}_a, \mathbf{S}_\xi) = \max\{\Upsilon_{ij}(\mathbf{S}_a, \mathbf{S}_\xi), i, j \in \mathcal{N}, i \neq j\}. \tag{7}$$

Note that Υ in some sense indicates the worst-case fairness for $\mathbf{S}_a$ and $\mathbf{S}_\xi$. Note also that an outcome is envy-free if $\Upsilon(\mathbf{S}_a, \mathbf{S}_\xi) = 1$.

In Fig. 4, we plot the envy ratio for an asymmetric outcome of the two-CN, two-channel selection stage game as a function of Δ, where Δ is the difference between the PU absent probabilities in the "preferred" channel and the "non-preferred channel". It can be seen in Fig. 4 that for $\Delta = 0$ the envy ratio is one, i.e., CNs are indifferent between the two sensing orders s_1 and s_2. As expected, the envy ratio increases as Δ increases. Moreover, it can also be seen that the envy ratio, for a fixed value of Δ, decreases for observations with false alarms. This seemingly counter-intuitive result can be explained as follows. For an asymmetric outcome when sensing observations are perfect, a CN can only transmit in its first step (if the channel is free). The CN that selects the sensing order s_1 with the "preferred" channel in its first step always transmits in that channel (if it is free) and the other CN that selects the sensing order with the "preferred" channel in its second step always transmits in the "non-preferred" channel (if it is free), as it finds the "preferred" channel busy. However, when the CNs have non-zero probabilities of false alarms, the CN with the "preferred" channel in its second step can be successful in finding the "preferred" channel free if the CN visits that channel and if it was PU-free but the other CN generated a false alarm in its first step. This reduces the envy ratio. Note, however, for observations with false alarms, the likelihood of finding a free channel for the CNs is reduced.

We then state the following result.

Proposition 4.1. *In the proposed sensing order selection game, the highest envy ratio is* $\frac{(1-\theta_1)}{(1-\theta_M)}$ *for the scenarios where* $N = M$ *and CNs have perfect observations.*

Proof. See [28, Proposition 4.1]. □

An asymmetric action profile corresponds to orthogonal sensing orders, i.e., each CN picks a different action. When $N \leq M$ there are M^N total possible outcomes and out of these total outcomes there are $\frac{M!}{(M-N)!}$ asymmetric outcomes. For efficient channel utilization, we consider the scenarios where the N CNs utilize the N top rows of Latin Square Φ for the selection of sensing orders. This is reasonable as the channel indices $1, 2, \ldots, M$ 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 PUs are less likely to be present. Note that for $N = M$, the entire matrix Φ of sensing orders is utilized by a CN for the selection of sensing orders.

5 Win-Shift Lose-Randomize (WSLR) Strategy for the Proposed Game, its Fairness and Complexity

5.1 WSLR Strategy

In this section, we propose an adaptive WSLR strategy for the channel sensing/access order selection game with imperfect observations, where adaptations are in the autonomous choices by the CNs, of the channel sensing order.

The WSLR strategy is described in Fig. 5a. The core idea of the WSLR strategy is as follows:

- The WSLR strategy utilizes adaptive randomization based on feedback for the CNs to autonomously arrive not only at orthogonal sensing orders but also at orthogonal time slots in a virtual frame of size V time slots, where $V = \lceil \frac{N}{M} \rceil$.

Win Shift Lose Randomize

1) Initialize $\mathbf{p} = [\frac{1}{N}, \frac{1}{N}, \cdots, \frac{1}{N}]$, if $N \leq M$, else initialize $\mathbf{p} = [\frac{1}{M}, \frac{1}{M}, \cdots, \frac{1}{M}]$, an N or M-element probability vector $\mathbf{p}$ (all components are nonnegative and add to 1), and initialize $\mathbf{b} = [\frac{1}{\lceil N/M \rceil}, \frac{1}{\lceil N/M \rceil}, \cdots, \frac{1}{\lceil N/M \rceil}]$ an $\lceil N/M \rceil$-element probability vector $\mathbf{b}$ (all components are nonnegative and add to 1).

2a) Toss a weighted coin to select a sensing order, with p_i being the probability of choosing sensing order i.

2b) Toss another weighted coin to select a time slot, with b_k being the probability of choosing time slot with kth index in the virtual frame of size $V = \lceil N/M \rceil$.

In the time slot k, sense the channels sequentially in the order given in the selected sensing order.

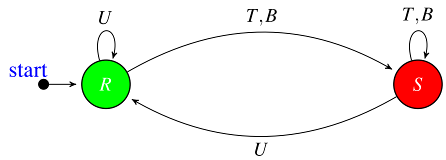
3) One of three possibilities occurs: a) Successful communication; b) CN finds all channels busy; and c) Unsuccessful communication.

a) or **b)** *Successful communication or CN finds all channels busy:* Using the current sensing order i, on a successful communication or a CN finds all channels busy, the CN updates $\mathbf{p}$ as $p_i = 1$, where $i = (i \bmod N) + 1$ and $p_j = 0, \forall j \neq i$, i.e., it shifts to the next sensing order to visit the channels in the next slot (**Switching Phase**). It also updates $b_k = 1$ and $b_l = 0, \forall l \neq k$, i.e., it selects the same time slot with index k. The CN then returns to 2.

c) *Unsuccessful communication:*

When the CN transmits but it receives no ACK in the current slot using sensing order i then the CN returns to 1 (**Randomization Phase**).

(a)



(b)

Fig. 5. a) Win-shift, lose-randomize (WSLR) strategy; b) two-state automaton representation of the WSLR strategy.

Note that when $N \leq M$, then $V = 1$, i.e., each CN updates its sensing order selection in every time slot. Before delving into the analysis, let us first present the high-level intuition that underlies the results in the rest of this paper. Consider N autonomous selfish CNs which do not share any information among themselves and can only use their own feedback to infer the action of other CNs. Moreover, they cannot store the history of their past observations and outcomes. When these autonomous CNs have to search multiple potentially available channels with non-homogeneous rewards for spectrum opportunities, they face competition from one another to access the channels with higher rewards. The end result of this competition is reduced CN throughput due to collisions among CNs that transmit simultaneously in the same channel. We are interested in finding a solution that allows CNs to autonomously achieve the following objectives:

- To autonomously converge to those channel sensing orders that minimize the likelihood of collisions among the CNs. In Section 5.3, we provide a theoretical upper bound for convergence to those sensing order selections that minimize the likelihood of collision among the CNs.
- To ensure fairness among distributed CNs competing for heterogeneous potentially available channels, in Section 5.4, we will explain how the proposed strategy achieves fairness among the CNs.
- To discourage selfish deviations, some punishment mechanism must be devised. A selfish CN that selects a channel sensing order with higher rewards

TABLE 2
Average Number of Time Slots Required by the WSLR Strategy to Arrive at Orthogonal Sensing Orders for Different $N = M$ Scenarios

	$N = M = 2$	$N = M = 4$	$N = M = 6$	$N = M = 8$	$N = M = 10$
WSLR $E[TTO]$	1.7	9.1	26.9	111.8	400.2

in a time slot may prefer to again select that sensing order in the next time slots, and such deviations should be discouraged through a punishment mechanism. In the proposed strategy, this is achieved by triggering a switch to the randomization phase when an unsuccessful transmission is observed. In Section 6, for different scenarios, we show that the proposed strategy is an equilibrium against unilateral selfish deviations.

5.2 Complexity of the Proposed WSLR Strategy

The complexity of a strategy can be measured in many different ways. We consider the complexity of the WSLR strategy in two different ways.

5.2.1 Complexity in the Context of Repeated Games Played by the Automaton

In this paper, we consider strategies for CNs that can be executed by a finite state machine called an automaton.

Definition 2. *A machine (automaton) A_i for CN i in an infinitely repeated game is defined by $A_i = \langle Q_i, q_i^0, f_i, \tau_i \rangle$, where Q_i is a finite set of states, q_i^0 is the initial state, f_i is an output function that assigns a pure behavior to every state, and τ_i is a finite set of transitions that determine for each state and for each observed signal, to which state to go next. The set of observed signals by a CN is Ξ, i.e., $\Xi = \{$Unsuccessful transmission (U), Successful transmission (T), Channels found busy $(B)\}$.*

When repeated games are played by an automaton, the number of states of the machine is often used as a measure of complexity. In our work, the proposed strategy requires only one-round of memory, as it responds to the previous round of play, and hence can be implemented by an automaton with only two states. Such a two-state machine $A_i = \langle Q_i, q_i^0, f_i, \tau_i \rangle$ of CN i (illustrated in Fig. 5b) carries out the WSLR strategy in the context of the opportunistic channel selection game and is explained as follows: $Q_i = \{R, S\}$, $q_i^0 = R$; $f_i(R)$ corresponds to the following behavior: for the selection of sensing/channel access orders, perform uniformly distributed random selection, i.e.,

$$\mathbf{p} = [1/N, 1/N, \ldots, 1/N], \tag{8}$$

where p_i represents the probability of selecting the ith sensing/channel access order and $\sum_{i=1}^{N} p_i = 1$; $f_i(S)$ corresponds to the following behavior: shift to the next sensing order to visit the channels in the next time slot.

With $q_i^0 = R$ being the start state, the rules of transition are given by: 1) from state R and with observed signal $\xi_i = U$, go to state R; 2) from state R and with observed signal $\xi_i = T$ or $\xi_i = B$, go to state S; 3) from state S and with observed signal $\xi = T$ or $\xi = B$, go to state S; and 4) from state S and with observed signal $\xi = U$, go to state R.

Due to imperfect observations, the machine of CN i in the opportunistic channel access game commits an error when for any given state the incorrect signal is observed due to channel error, co-channel interference tolerance and false alarm. These errors will lead the machine of the CNs to erroneously move from the randomization state R to the shifting state S, or from the shifting state S to the randomization state R. For instance, a false alarm or the capture effect may have the effect of a CN thinking that the two CNs have selected different sensing orders when in fact they may have selected the same sensing order. Moreover, imperfect observations due to channel errors may have the effect of a CN thinking that the two CNs have selected the same sensing order when in fact they may have selected different sensing orders.

5.2.2 Complexity in Terms of Information Exchange Overhead

The only control information our method requires is: Initially, each CN obtains the statistics of primary users' duty cycles and the number of other active CNs through a spectrum access system and later there is no further control information exchange in the network. The proposed WSLR strategy involves little information exchange overhead and can be be executed by a two state machine. Moreover, it makes adaptive channel selection decisions based on simple feedback outcomes (such as occurrence of a successful transmission or a collision). These features along with the recent advances in the implementation of programmable wireless nodes (see [37], and references therein) makes it possible that the proposed strategy can be incorporated into wireless nodes that are built on flexible software defined platforms.

We next provide some results pertaining to convergence.

5.3 Convergence to Conflict-Free Channel Sensing/Access Orders

In Table 2 we provide average number of time slots (simulated) required by the WSLR strategy to arrive at orthogonal sensing orders for different scenarios.

Using the WSLR strategy, the stochastic process of channel sensing/access order selection can be modeled as a finite-state absorbing Markov chain with N states, where N is the number of CNs. For instance, when $N = 4$ there are four states in the Markov chain, a state (4) means that all four CNs randomly select a channel sensing order in a given round, a state $(3,1)$ means that three CNs randomly select while one CN does not perform random selection, a state $(2,1,1)$ means that two CNs randomly select while two CNs do not perform random selection, and state $(1,1,1,1)$ is an absorbing state, as no CN performs random selection. Using the above absorbing Markov chain model one can theoretically calculate the number of time slots required for convergence for a particular given N, but for any N a closed form expression is difficult to obtain. Although it is difficult to derive a closed-form expression for the average number of time slots required for convergence, we provide a theoretical upper bound for convergence to those sensing order selections that minimize the likelihood of collision among the CNs.

Proposition 5.1. *For $N \leq M$, the expected time-to-orthogonalize (E[TTO]) using the WSLR strategy is $E[\mathrm{TTO}] \leq \frac{N^N}{N!}$ time slots.*

Proof. See Appendix B, available in the online supplemental material. □

For the case of imperfect observations (due to channel error) of feedback, we have performed comprehensive numerical analysis in Section 7, and have shown that when CNs have false alarms and channel errors the WSLR strategy still outperforms other autonomous strategies.

5.4 Fairness

The proposed strategy ensures fairness among the CNs by allowing them to autonomously share across time non-homogeneous channel rewards. When all CNs play WSLR, then in the steady state, all CNs will keep switching among N sensing orders one by one. Under steady state, in a given round t, when CN i selects sensing order s_1 it gets a reward of $g_i(s_1, (s_2, s_3, \ldots, s_N))$, when CN j selects s_2, it gets a reward of $g_j(s_2, (s_1, s_3, \ldots, s_N))$, and so on. Then in round $t+1$, CN i selects sensing order s_2, and it gets a reward of $g_i(s_2, (s_1, s_3, \ldots, s_N))$, CN j selects s_3, and it get a reward of $g_j(s_3, (s_1, s_2, \ldots, s_N))$, and so on. This process of shifting to the next sensing order is repeated and by sharing the N sensing orders across time (rounds), the proposed strategy equalizes everyone's rewards and hence it ensures envy-freeness. Numerical results related to the envy-freeness of the proposed strategy are discussed in detail in the performance evaluation section (Section 7).

6 Analysis of the Proposed WSLR Strategy and Repeated Game

In this section, we present analytical results on the performance of the proposed channel selection strategy for sequential sensing. In Section 7, we also present results for the scenarios where the CNs can sense only one channel in a time slot.

6.1 Two-CN Two-Channel Scenario

When both CNs implement their strategies by automata A_1 and A_2, the pair of automata (A_1, A_2) forms a system which can be analyzed with a finite Markov chain. This allows us to use the framework of Markov chains with rewards to calculate analytically the expected payoff of a CN [38].

The state of the Markov chain is characterized by the tuple $\omega = (\mathbf{s}, \mathbf{q}^{\xi})$, where $\mathbf{s} = (s_i, s_{-i})$ is an action profile of the two CNs and $\mathbf{q}^{\xi} = (q_i^{\xi}, q_{-i}^{\xi})$ corresponds to the signals observed. For instance, the situation where the system is in state $\big((s_1, s_2), (U, U)\big)$ corresponds to the scenario where the two CNs have selected different sensing orders, the automata of both CNs are in the shift state and they both observe unsuccessful transmissions signal (U). We represent the transition probabilities by $P_{\omega\omega'}$ and the state space of the Markov chain is represented by Ω (see Appendix C, available in the online supplemental material, for the details of the calculation of expected payoffs using Markov chains with rewards).

We show in [29] that the Markov chain above is an ergodic unichain. The steady state reward per step for a CN i is then independent of the starting state and is given by

$$\bar{G}_i(A_i, A_{-i}) = \sum_{j \in \Omega} \delta_j \hat{g}_{j,i}, \tag{9}$$

where δ_j is the steady state probability of the jth state and $\hat{g}_{j,i}$ is the reward associated with the jth state for an individual CN i.

When the two CNs with observations utilize the adaptive WSLR strategy there are $|\Omega| = 12$ states of the Markov chain, and when the CNs have imperfect observations there are $|\Omega| = 36$ states of the Markov chain.

Proposition 6.1. *The WSLR strategy for the two-CN two-channel scenario (when adopted by both CNs that utilize automata with no more than two states and have perfect observations) is a Nash Equilibrium.*

Proof. See Appendix D, available in the online supplemental material. □

We next derive the conditions under which the WSLR leads to a Nash Equilibrium, in the case of imperfect observations.

Proposition 6.2. *The WSLR strategy for the two-CN two-channel scenario (when adopted by both CNs that utilize automata with no more than two states and have imperfect observations) is a Nash Equilibrium for $P_{fa}, \pi, \sigma \in [0, 0.2]$.*

Proof. See Appendix E, available in the online supplemental material. □

6.2 N-CN and M-Channel Scenario, $N \leq M$

The steady-state payoff per time slot (when all autonomous CNs with perfect observations play A^{WSLR}, always WSLR) for a CN is at least $\sum_{i=1}^{N}(1-\theta_i)/N$, when $N < M$, and exactly $\sum_{i=1}^{N}(1-\theta_i)/N$ for $N = M$. The reason is as follows. Consider $N < M$ autonomous CNs that can sense only one channel within the duration of a slot, i.e., $k = 1$. It is easy to see that in steady state, the CNs will keep switching among the top N sensing orders of matrix Φ, and the steady state reward per CN per time slot is $\sum_{i=1}^{N}(1-\theta_i)/N$. Now consider the case when each CN can sense more than one channel sequentially within the duration of a slot, i.e., $k > 1$. Note that in the steady state, the probability of success in the first step of any CN is not affected by the other competing CNs. However, in the later steps a CN can obtain a channel if it visits a PU-free channel and no other CN has already found that PU-free channel (in the previous steps).

Hence, the expected reward of a CN is at least $\sum_{i=1}^{N}(1-\theta_i)/N$. For $N=M$, in steady state, each CN can only find a free channel in the first sensing step, hence the steady state reward per time slot for a CN is exactly $\sum_{i=1}^{N}(1-\theta_i)/N$.

Deriving the proof for equilibrium when $N \leq M$, where $N > 2$, is challenging due to the combinatorial explosion in the number of ways that N CNs can find channels free from PUs and other CNs, and also the number of ways N CNs can generate false alarms, experience channel errors and the capture effect. To keep the analysis tractable, we next provide proofs that an equilibrium exists for $N \leq M$ CNs with perfect observations for the scenarios where: 1) all potentially available channels have the same expected payoff values; and 2) different potentially available channels offer different expected payoff values. Moreover, in Section 7, through extensive simulations we will analyze the performance of the WSLR strategy for CNs with imperfect observations using a synthetic PU channel occupancy model and also real PU spectrum occupancy data.

Proposition 6.3. *The WSLR strategy for N autonomous CNs and M potentially available channels, where $N \leq M$, when adopted by all CNs that utilize the automata with no more than two states, is an equilibrium for potentially available channels with homogeneous rewards.*

Proof. See Appendix F, available in the online supplemental material. □

To show the existence of an equilibrium when potentially available channels have different reward values, we first provide an upper bound on the reward of a CN i that deviates from the WSLR strategy while all the other $(N-1)$ CNs follow the WSLR strategy.

Proposition 6.4. *In the proposed game when a CN, say i, deviates from A^{WSLR} by playing A_i^{WB} (Weighted Best strategy), i.e., by selecting $s_1 = (1,2,3,4,\ldots,M)$, the preferred sensing order, with probability $q \in (0.5, 1]$ and s_j with probability $\frac{(1-q)}{(N-1)}$, $\forall$ $s_j \in \mathcal{S}$, $s_j \neq s_1$, while all the other CNs play A^{WSLR}, then its expected reward per time slot is strictly less than $(1-\frac{1}{N})^{N-1}$.*

Proof. See Appendix G, available in the online supplemental material. □

Next, we use Proposition 6.4 to show that playing A^{WSLR} leads to an Equilibrium.

Proposition 6.5. *When $\sum_{i=1}^{N}(1-\theta_i)/N \geq (1-\frac{1}{N})^{N-1}$ then for N CNs and M potentially available channels with non-homogeneous reward values, where $N \leq M$, playing A^{WSLR}, when adopted by all CNs that play automata with no more than two states, is an equilibrium against the* Weighted Best *deviation.*

Proof. See Appendix H, available in the online supplemental material. □

For example, when $N = 4$ CNs, the maximum expected reward per time slot that the deviating CN i can obtain is 0.42 and it decreases with increasing N and M. Hence, for $\theta_i \in [0, 0.58]$, $i \in \mathbf{M}$, playing A^{WSLR}, when adopted by all CNs that use automata with no more than two states, is an equilibrium against the *Weighted Best* deviation. We note that in practice, this is indeed a reasonable condition, as generally, it may not be efficient in terms of CN throughput to allow opportunistic spectrum access in spectrum bands where the probability of PU being active is high. Moreover, the CNs may not be able to sense all the channels due to limited sensing and hardware capability. In order to increase channel utilization efficiency, CNs should focus on utilizing channels with low to medium PU activity [39].

7 Performance Evaluation

7.1 Simulation Results

Through extensive simulations, we evaluate and compare the performance of the WSLR strategy for perfect and imperfect observations using both synthetic and real PU spectrum data.

Using simulations we evaluate the performance of the proposed strategy in terms of: 1) total average payoff (total average number of successful transmissions) per time slot in the CN network ($\sum_{i=1}^{N} G_i$); 2) expected payoff of a CN i per time slot (G_i); 3) the highest value of the envy ratio Υ between a pair of CNs; 4) average total unsuccessful transmissions in the CN network; 5) the effect of varying the number of available sensing steps; and 6) the probability of finding a channel free in the first sensing step (given that the CN is successful). Our aim is also to compare the performance of the WSLR strategy against: 1) when all CNs utilize random selection of sensing orders (*random strategy*); 2) when a centralized entity orthogonally allocates and shifts sensing orders for each CN (*centralized*); and 3) an autonomous CN i considers deviating from the WSLR strategy while all other CNs follow the strategy. The studied deviations by the CN i are: a) always select the most preferred sensing order $s_1 = (1,2,\ldots,M)$, *always best (AB)*; and b) always select s_1 with probability $q = 0.75$ and s_2 with probability $(1-q)$, *weighted best (WB) deviation*.

Figs. 6a, 6b, and 6c evaluate the total expected payoff per time slot in the network, total expected collisions per time slot in the network and expected payoff of a CN per time slot under different scenarios. From Fig. 6a we can see that using the WSLR strategy the network achieves almost the same total expected payoff as achieved by a centralized allocation. In Fig. 6b it can be seen that the total number of collisions per time slot in the network for the WSLR strategy are significantly fewer when compared to the other strategies (apart from centralized allocation). Moreover, from Figs. 6a, 6b, and6c we can see that the WSLR strategy performs significantly better as compared to when a CN considers deviating while all the other CNs follow the WSLR strategy, or when all N CNs in the network utilize random selection of sensing orders.

Next, in Figs. 7a and 7b we explore the impact of imperfect observations on the expected payoff per time slot of a CN i as a function of $N = M$ total CNs in the network. From the two figures we can see that the WSLR strategy still achieves the highest expected payoff per time slot for the CN i as compared to the other distributed strategies. Note that in Figs. 7a and 7b the loss in the expected reward of the CN i as $N = M$ increases is due to the non-homogeneity in channel availability statistics, i.e., due to primary user

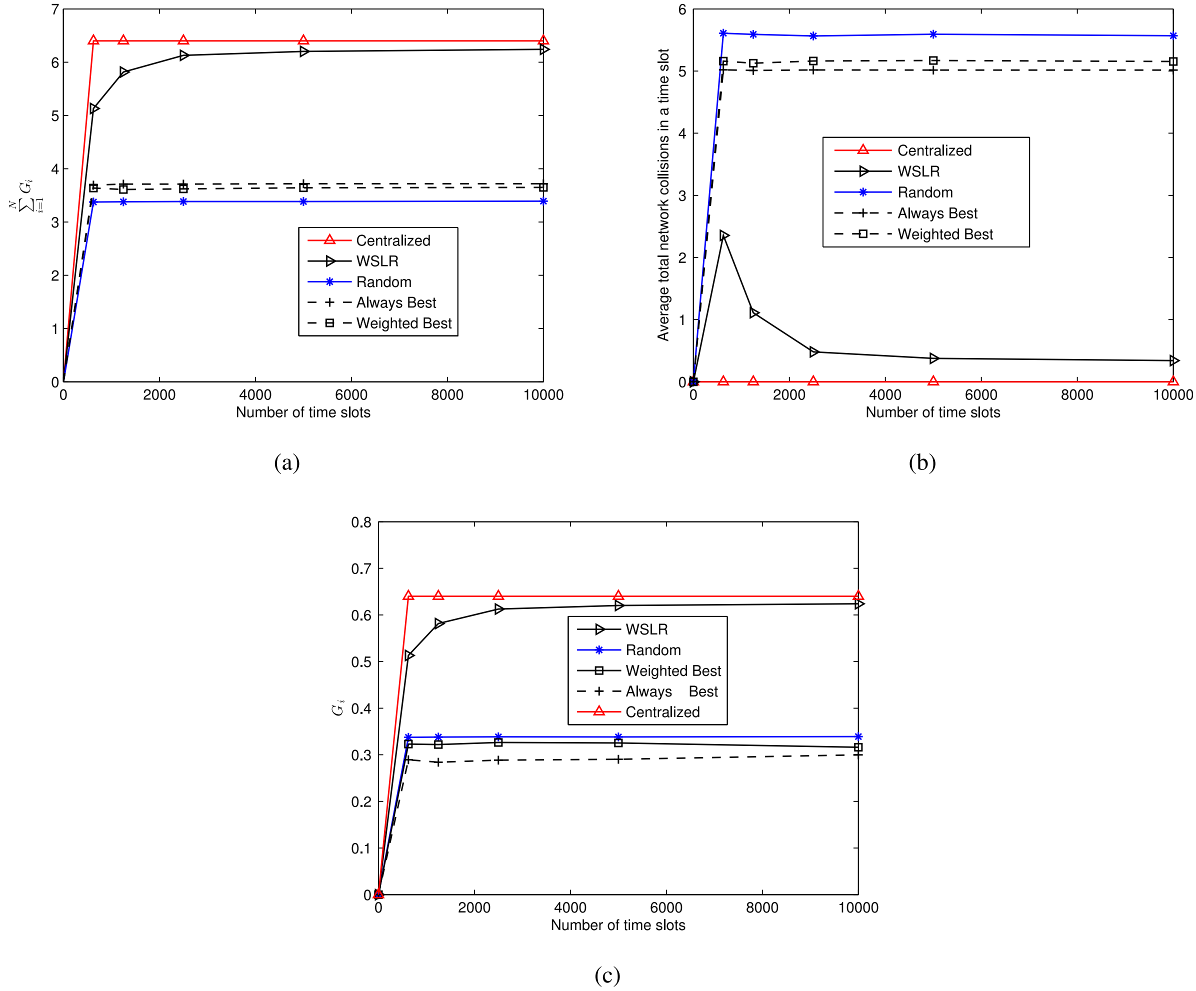


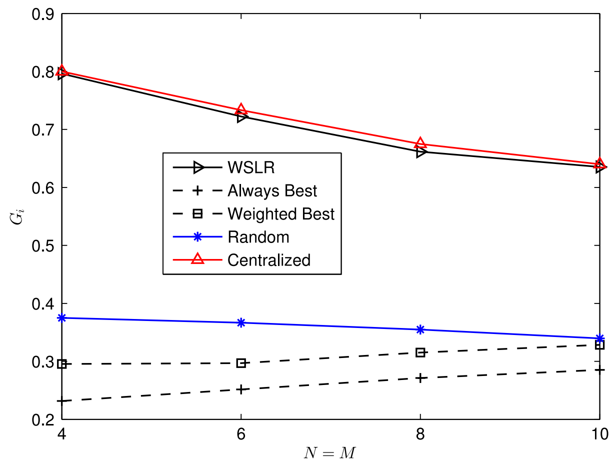
Fig. 6. For $N = M = 10$ CNs with perfect observations, a) Total expected reward per time slot in the network $(\sum_{i=1}^{N} G_i)$ as a function of time slots for different scenarios; b) total expected unsuccessful transmissions in the network as a function of time slots; and c) expected payoff per time slot for a CN i (G_i) as a function of time slots. In all considered scenarios $\Theta = (0.1, 0.2, 0.2, 0.3, 0.3, 0.5, 0.5, 0.5, 0.5, 0.5)$ represents the primary user duty cycle statistics vector for channels 1 to M, respectively.

duty cycle statistics vector $\Theta = (0.1, 0.2, 0.2, 0.3, 0.3, 0.5, 0.5, 0.5, 0.5, 0.5)$. The availability probabilities of the first 5 channels are at least 70 percent and the availability probabilities of the last five channels are 50 percent. Hence, as $N = M$ increases, the expected payoff of the CN i decreases, as with the increasing number of CNs the number of potentially available channels also increases but with high probability of a PU being present. Moreover, the loss in the expected payoff of the CN i with increasing $N = M$ is also due to channel errors. When $\pi > 0$, in a given time slot if all the CNs arrive at orthogonal sensing orders, then an unsuccessful transmission (due to channel error) in any later time slot will lead a CN to erroneously move from shifting to randomization phase, in which case orthogonality may be lost, leading to reduced reward values. However, Figs. 7a and 7b also show that for imperfect observations our proposed strategy enables the CNs to increase their payoff as compared to other distributed strategies. It also shows that when observation errors are not significant the proposed strategy performs similarly to the centralized strategy.

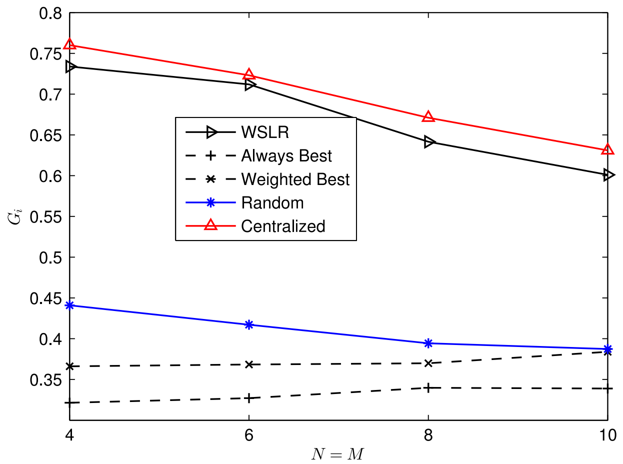
In Table 3, we evaluate $P^*_{s,1}$, i.e., the probability of finding a channel free in the first step (given that the channel is free), for the scenarios where the CNs have perfect sensing and where CNs have false alarms in their observations. It can be seen that using the WSLR strategy $P^*_{s,1}$ is between 0.96 to 0.98. This reduces the overhead of multiple sensing steps incurred by a CN. The reduced number of sensing steps required to find a channel free in turn increases the throughput per time slot of a CN.

7.2 Performance Evaluation Using Real Spectrum Occupancy Data

In Table 4, we evaluate the effectiveness of our proposed strategy for perfect and imperfect observations by testing it with real spectrum occupancy data collected in the DECT bands. It can be seen that in all scenarios the WSLR strategy performs equally well as centralized orthogonal allocation of sensing orders for perfect observations and it outperforms all other strategies. Due to imperfect observations, there is some degradation in performance as compared to the centralized orthogonal allocation of sensing orders but the proposed strategy still outperforms the other strategies. Moreover, it can also be seen that the WSLR strategy also ensures envy-freeness among the competing CNs. Table 4 also shows that when the random selection of sensing orders (rand) is



(a) The vector of observation error probabilities $\mathbf{e}$ is set to $\mathbf{e} = (P_{fa}, \sigma, \pi) = (0.1, 0, 0)$.



(b) The vector of observation error probabilities $\mathbf{e}$ is set to $\mathbf{e} = (P_{fa}, \sigma, \pi) = (0.0, 0.05, 0.05)$.

Fig. 7. Expected payoff per time slot of the CN i (G_i) as a function of $N = M$ CNs for different scenarios. $\Theta = (0.1, 0.2, 0.2, 0.3, 0.3, 0.5, 0.5, 0.5, 0.5, 0.5)$ represents the primary user duty cycle vector for channels 1 to M.

adopted by all CNs, it ensures fairness; however, it significantly reduces the total average payoff per time slot in the CN network. When a CN considers deviating, while all other CNs follow the WSLR strategy (see always best and weighted best in Table 4), it can be seen that there is no incentive to deviate from the WSLR strategy. Moreover, when a CN deviates, although the total average payoff per time slot in the CN network is higher when compared to the random selection of sensing orders, a deviating CN significantly reduces the total expected payoff per time slot as compared to when all the CNs adopt the WSLR strategy. This is consistent with the results shown for synthetic data.

TABLE 4
Highest Average Envy Ratio Υ Between A Pair of CNs in the Network, and Average Total Payoff in the Network Per Time Slot $\sum_{i=1}^{8} G_i$ as A Function of $N = M = 8$ for Different Scenarios and Strategies

	$P_{fa} = \pi = \sigma = 0$		$P_{fa} = 0.1, \pi = \sigma = 0$		$P_{fa} = 0, \pi = \sigma = 0.05$	
	Υ	$\sum_{i=1}^{8} G_i$	Υ	$\sum_{i}^{8} G_i$	Υ	$\sum_{i=1}^{8} G_i$
Centralized	$\frac{0.37}{0.37} = 1$	2.96	$\frac{0.37}{0.37} = 1$	2.928	$\frac{0.35}{0.35} = 1$	2.85
WSLR	$\frac{0.369}{0.369} = 1$	2.95	$\frac{0.36}{0.36} = 1$	2.912	$\frac{0.31}{0.31} = 1$	2.4
Random	$\frac{0.20}{0.20} = 1$	1.6	$\frac{0.23}{0.23} = 1$	1.852	$\frac{0.20}{0.20} = 1$	1.6
Always Best	$\frac{0.236}{0.162} = 1.46$	1.83	$\frac{0.26}{0.19} = 1.38$	1.94	$\frac{0.23}{0.16} = 1.4$	1.65
Weighted Best	$\frac{0.236}{0.163} = 1.44$	1.81	$\frac{0.25}{0.18} = 1.4$	1.96	$\frac{0.22}{0.15} = 1.5$	1.7

Real spectrum occupancy data collected in the DECT bands by RWTH Aachen is utilized for channels 1 to M.

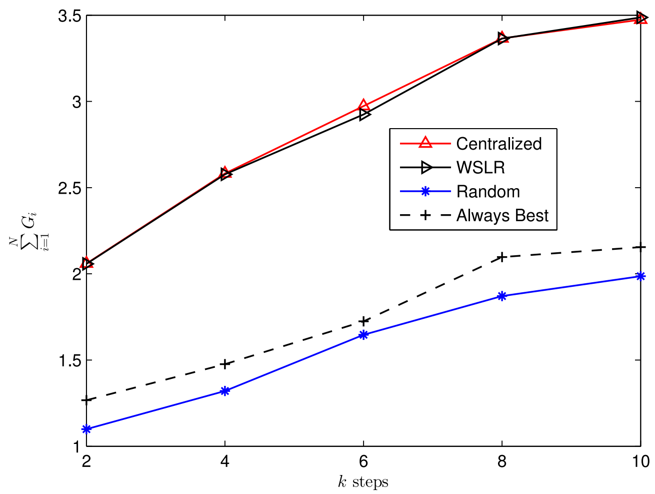


Fig. 8. Total expected payoff per time slot $(\sum_{i=1}^{N} G_i)$ as a function of the number of sensing steps k, with $N = 5$ CNs, $M = 10$ channels (sequential channel sensing).

Unlike previous scenarios where $N = M$ CNs were active in the network, next we consider a scenario where $N < M$. Using the real spectrum occupancy data collected in the DECT bands by RWTH Aachen, in Fig. 8, we consider sequential channel sensing and evaluate the effect of varying the number of sensing steps on the performance of the different strategies in terms of total expected payoff in the network per time slot. It can be seen in Fig. 8 that when all the CNs utilize the WSLR strategy then the expected payoff per time slot of a CN increases as the number of sensing steps increases. However, it can also be seen in Fig. 8 that in the presence of a deviating CN, when all other CNs follow the WSLR strategy, there is little 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 CNs utilize the

TABLE 3
Probability of Finding a Channel Free in First Step (Given that the Channel is Free) when $N = M = 10$

	$P_{fa} = 0$				$P_{fa} = 0.1$			
	WSLR	Random	Weighted Best	Fixed Best	WSLR	Random	Weighted Best	Fixed Best
$P_{s,1}^{*}$	0.9895	0.7302	0.7476	0.7352	0.962	0.7358	0.7483	0.7338

TABLE 5
Average Total Payoff in the Network Per Time Slot $\sum_{i=1}^{20} G_i$ for $N = 20$ CNs with Perfect Observations and $M = 10$ Channels

Centralized Orthogonal	WSLR	Random
10	9.78	5.65

Average number of channels free from PU activity per time slot is 10.

WSLR strategy, it becomes more likely for a CN to find free channels in the later sensing steps as the WSLR strategy allows them to arrive at conflict-free allocations. When a CN deviates, while all other CNs stay on the WSLR strategy, the CN can either find a free channel in its initial sensing steps when the CN is the sole radio following this sensing order, or it may collide with the other CNs when some other CN also selects the same sensing order and fails to find a free channel during that time slot. Hence, further increasing the number of sensing steps provides little gain.

For $N = 20$ CNs and $M = 10$ channels, we evaluate and compare the performance of the WSLR strategy with random selection and also with the centralized strategy in Table 5. It can be seen that the WSLR strategy performs similarly to the centralized orthogonal strategy.

8 Conclusions

We have studied the problem of coexistence among multiple autonomous cognitive nodes which compete for a common pool of potentially available channels. We have considered the real spectrum occupancy data collected at RWTH Aachen and found that spectrum resources can be non-homogeneous in terms of primary user occupancy. The non-homogeneity in resources leads to conflict in payoff distribution among autonomous CNs in the network. To address this challenge, we have designed an adaptive win-shift lose-randomize strategy. To analyze the impact of imperfect observations and/or selfish deviations of a CN on the stability of the proposed WSLR strategy, we have utilized the framework of repeated games with imperfect observations and limited memory. We have shown that the proposed strategy maximizes the total average number of successful transmissions in the network, ensures fairness by allowing the autonomous CNs to engage in inter-temporal sharing of the non-homogeneous rewards from cooperation, and allows the CNs to find a free channel quickly as compared to the other strategies. Moreover, using real PU spectrum occupancy data, we have conducted extensive simulations and have explored the effects of false alarm, channel error, and co-channel interference tolerance on the performance of our proposed strategy. We have found that results from the real primary spectrum occupancy data are consistent with the results shown for synthetic data.

Acknowledgments

This work was funded by Academy of Finland, National Science Foundation, USA, Science Foundation Ireland, Research Council of the University of Oulu, and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). The material in this paper was presented in part at the IEEE Global Communications Conference (GLOBECOM'13), 2013. [Online]. Available: http://www.ee.oulu.fi/ zaheer/Globecom2013.pdf. Z. Khan is the corresponding author.

References

[1] E. E. Schmidt, and C. Mundie, (2012, Jul.). Realizing the full potential of government-held spectrum to spur economic growth.. [Online]. Available: http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast_spectrum_report_final_july_20_2012.pdf

[2] J. Melvin, (2012, Aug.). US regulators ok T-mobile testing of shared use of airwaves.. [Online]. Available: http://www.reuters.com/article/2012/08/15/usa-spectrum-sharing-idUSL2E8JFD6M20120815

[3] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 32–39, Apr. 2008.

[4] M. Wellens and P. Mähönen, "Lessons learned from an extensive spectrum occupancy measurement campaign and a stochastic duty cycle model," *Mobile Netw. App.*, vol. 15, no. 3, pp. 461–474, 2010.

[5] A. Rubinstein, *Modeling Bounded Rationality*, 2nd ed. Cambridge, MA, USA: MIT Press, 1998.

[6] J. Romero, (2011). Finite automata in undiscounted repeated games with private monitoring. [Online]. Available: http://EconPapers.repec.org/RePEc:pur:prukra:1260

[7] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 731–745, Apr. 2011.

[8] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5667–5681, Nov. 2010.

[9] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 785–796, Feb. 2008.

[10] K. Liu, Q. Zhao, and B. Krishnamachari, "Distributed learning under imperfect sensing in cognitive radio networks," in *Proc. Conf. Record 44th Asilomar Conf. Signals, Syst. Comput.*, 2010, pp. 671–675.

[11] Y. Gai, B. Krishnamachari, and R. Jain, "Learning multiuser channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation," in *Proc. New Frontiers Dyn. Spectrum*, 2010, pp. 1–9.

[12] 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.

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

[14] N. B. Chang and M. Liu, "Competitive analysis of opportunistic spectrum access strategies," in *Proc. IEEE Int. Conf. Comput. Commun.*, Apr. 2008, pp. 1535–1542.

[15] Z. Khan, J. Lehtomaki, L. DaSilva, and M. Latva-aho, "Autonomous sensing order selection strategies exploiting channel access information," *IEEE Trans. Mobile Comput.*, vol. 12, no. 2, pp. 274–288, Feb. 2013.

[16] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.

[17] W. Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Repeated open spectrum sharing game with cheat-proof strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1922–1933, Apr. 2009.

[18] X. Yuanzhang and M. van der Schaar, "Repeated resource sharing among selfish players with imperfect binary feedback," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2012, pp. 452–459.

[19] X. Yuanzhang and M. van der Schaar, "Nonstationary resource sharing with imperfect binary feedback: An optimal design framework for cost minimization," in *Proc. 51st Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 932–939.

[20] X. Y. Xiao and M. van der Schaar, "Dynamic spectrum sharing among repeatedly interacting selfish users with imperfect monitoring," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1890–1899, Nov. 2012.

[21] K. Bian, J. M. Park, and B. Gao, *Cognitive Radio Networks: Medium Access Control for Coexistence of Wireless Systems*. New York, NY, USA: Springer, 2014.

[22] W. Wenjing, M. Chatterjee, and K. Kwiat, "Cooperation in wireless networks with unreliable channels," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2808–2817, Oct. 2011.

[23] N. See-Kee and W. Seah, "Game-theoretic model for collaborative protocols in selfish, tariff-free, multihop wireless networks," in *Proc. 27th IEEE Conf. Comput. Commun.*, Apr. 2008, pp. 216–220.

[24] V. R. Srivastava, "Behavior-based incentives for node cooperation in wireless ad hoc networks," Ph.D. dissertation, Virginia Polytechnic Inst. State Univ. (Virginia Tech), Arlington, TX, 2008.

[25] Z. Biling, C. Yan, and K. Liu, "An indirect-reciprocity reputation game for cooperation in dynamic spectrum access networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4328–4341, Dec. 2012.

[26] W. Chih-Yu, C. Yan, and K. Liu, "Sequential chinese restaurant game," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.

[27] J. Chunxiao, C. Yan, Y. Yu-Han, W. Chih-Yu, and K. Liu, "Dynamic chinese restaurant game in cognitive radio networks," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2013, pp. 962–970.

[28] Z. Biling, C. Yan, W. Chih-Yu, and K. Liu, "Learning and decision making with negative externality for opportunistic spectrum access," in *Proc. IEEE Global Commun. Conf.*, Dec. 2012, pp. 1404–1409.

[29] Z. Khan, J. J. Lehtomaki, L. A. DaSilva, M. Latva-aho, and M. Juntti, "Adaptation in a channel access game with private monitoring," in *Proc. IEEE Global Commun. Conf.*, 2013, pp. 3157–3163.

[30] R. Fan and H. Jiang, "Optimal multi-channel cooperative sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1128–1138, Mar. 2010.

[31] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE Int. Dyn. Spectrum Access Netw.*, 2005, pp. 131–136.

[32] E. Peh and Y.-C. Liang, "Optimization for cooperative sensing in cognitive radio networks," in *Proc. IEEE Int. Wireless Commun. Netw. Conf.*, 2007, pp. 27–32.

[33] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[34] K. Whitehouse, A. Woo, F. Jiang, J. Polastre, and D. Culler, "Exploiting the capture effect for collision detection and recovery," in *Proc. 2nd IEEE Workshop Embedded Netw. Sens.*, May 2005, pp. 45–52.

[35] M. J. Osbourne, *An Introduction to Game Theory*. London, U.K.: Oxford Univ. Press, 2004.

[36] R. J. Lipton, E. Markakis, E. Mossel, and A. Saberi, "On approximately fair allocations of indivisible goods," in *Proc. 5th ACM Conf. Electron. Commerce*, 2004, pp. 125–131.

[37] P. D. Francesco, S. McGettrick, U. Anyanwu, J. O'Sullivan, A. MacKenzie, and L. DaSilva, "A split MAC approach for SDR platforms," *IEEE Trans. Comput.*, vol. 64, no. 4, pp. 912–924, Apr. 2015.

[38] R. G. Gallager, *Discrete Stochastic Processes*. Boston, MA, USA: Kluwer, 2001.

[39] C. Tekin, S. Hong, and W. Stark, "Enhancing cognitive radio dynamic spectrum sensing through adaptive learning," in *Proc. IEEE Military Commun. Conf.*, 2009, pp. 1–7.

Zaheer Khan received the Dr.Sc degree in telecommunications from the University of Oulu, Oulu, Finland, in 2011, and the MSc degree in electrical engineering from the University College Borås, Sweden, in 2007. He is the recipient of the Marie Curie fellowship for the year 2007 to 2008. His research interests include application of game theory to model distributed wireless networks, cognitive and cooperative communications, and wireless signal design. Specifically, he has been working on the design of distributed and adaptive resource allocation strategies for cognitive radio networks.

Janne J. Lehtomäki graduated with an MSc and PhD degrees in telecommunications from the University of Oulu, Finland, in 1999 and 2005, respectively. Currently, he is a senior research fellow at the University of Oulu, Centre for Wireless Communications (CWC). His research interests include nanonetworks, spectrum measurements, energy detection, and cognitive radio networks. He coauthored the paper receiving the Best Paper Award in IEEE WCNC 2012.

Luiz A. DaSilva holds the stokes professorship in telecommunications in the Department of Electronic and Electrical Engineering at Trinity College Dublin, where he is a co-principal investigator of CTVR, The Telecommunications Research Centre in Ireland. He is also a professor in the Bradley Department of Electrical and Computer Engineering at Virginia Tech. His research focuses on distributed and adaptive resource management in wireless networks, and in particular cognitive radio networks, and the application of game theory to wireless networks. He is currently a principal investigator on research projects funded by the US National Science Foundation, the Science Foundation Ireland, and the European Commission under Framework Programme 7. He has authored more than 150 peer-reviewed journal and conference papers and two books on wireless communications.

Ekram Hossain (F'15) received the PhD degree in electrical engineering from the University of Victoria, Canada, in 2001. He is a professor (since March 2010) in the Department of Electrical and Computer Engineering at University of Manitoba, Winnipeg, Canada. His current research interests include design, analysis, and optimization of wireless/mobile communications networks, cognitive radio systems, and network economics. He serves as the editor-in-chief for the *IEEE Communications Surveys and Tutorials* and an editor for *IEEE Wireless Communications*. Also, he is a member of the IEEE Press Editorial Board. Previously, he served as the area editor for the *IEEE Transactions on Wireless Communications* in the area of "Resource Management and Multiple Access" from 2009 to 2011, an editor for the *IEEE Transactions on Mobile Computing* from 2007 to 2012, and an editor for the *IEEE Journal on Selected Areas in Communications-Cognitive Radio Series* from 2011 to 2014. He has received several research awards including the University of Manitoba Merit Award in 2010 and 2014 (for Research and Scholarly Activities), the 2011 IEEE Communications Society Fred Ellersick Prize Paper Award, and the IEEE Wireless Communications and Networking Conference 2012 (WCNC'12) Best Paper Award. He is a distinguished lecturer of the IEEE Communications Society (2012-2015). He is a registered professional engineer in the province of Manitoba, Canada. He is a fellow of the IEEE.

Matti Latva-Aho received the MSc (EE), Lic. Tech. and Dr. Tech degrees from the University of Oulu, Finland in 1992, 1996, and 1998, respectively. From 1992 to 1993, he was a research engineer at Nokia Mobile Phones in the CDMA systems research group. During the years 1994 to 1998, he was a research scientist at the Telecommunication Laboratory and Centre for Wireless Communications (CWC) at the University of Oulu.