A game-theoretic and stochastic survivability mechanism against induced attacks in Cognitive Radio Networks

Saad Mneimneh a, Suman Bhunia b,∗, Felisa Vázquez-Abad a, Shamik Sengupta b

a Department of Computer Science, Hunter College, City University of New York, 10065, USA
b Department of Computer Science and Engineering, University of Nevada, Reno, 89557, USA

ABSTRACT

Cognitive Radio Networks (CRNs) are envisioned to provide a solution to the scarcity of the available frequency spectrum. It allows unlicensed secondary users (SUs) to use spectrum bands that are not occupied by licensed primary users (PUs) in an opportunistic manner. This dynamic manner of spectrum access gives rise to vulnerabilities that are unique to CRNs. In the battle over the available spectrum, SUs do not have any means of identifying whether disruption sensed on a band is intentional or unintentional. This problem is further intensified in the case of heterogeneous spectrum, where different bands provide different utilities. A smart malicious agent can use this vulnerability to temporarily disrupt transmissions on certain bands and induce their unavailability on SUs. The motivation for such disruption-induced attacks can be either monopolism, i.e. to capture as much spectrum as possible and make other SUs starve, or denial of service by intentional disruption of other SUs’ communications. This paper proposes an adaptive strategy for robust dynamic spectrum access in the event of induced attacks. Assuming rational players, and considering the notion of channel utility, the optimal strategy is established by modeling such scenarios as zero-sum games that lead to Nash equilibrium. Thereafter, the case of non-stationary channel utilities is investigated, where utilities are subject to abrupt changes due to fluctuations in channel characteristics, as well as arrival and departure of PUs. Through concurrent estimation, learning, and optimal play, it is shown that the proposed mechanism performs robustly even in such dynamic environments. Comparison of the proposed mechanism to other reasonable benchmark strategies in simulation confirms that this mechanism significantly enhances the performance of CRNs.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The conventional fixed spectrum assignment policy has resulted in suboptimal use of spectrum resource leading to over-utilization in some bands and under-utilization in others [1,2]. This observation has led to the recent spectrum policy...
reforms (Dynamic Spectrum Access) by the US Federal Communications Commission (FCC). The goal of dynamic spectrum access (DSA) is expected to be achieved via the recently proposed concept of the cognitive radio (CR) [3]. In contrast to the legacy radios, CRs are envisioned to intelligently adjust their transmission/reception parameters themselves and find the best available spectrum bands to use [4–9]. Along this line, IEEE has setup a working group (IEEE 802.22) to develop the standard for CR based Wireless Regional Area Networks (WRAN). The standard specifies the physical and medium access control layer of the fixed and portable point-to-multipoint WRAN operating over the spectrum allocated to TV broadcasting services. The most important regulatory aspect is that cognitive radios must dynamically identify and opportunistically access unused spectrum bands, but must not interfere with the primary users (licensed holders) operating in licensed bands and, therefore, switch to a different band promptly upon sensing the primary user.

The challenge of survivability under adversarial conditions: With the advent of “smartsness” and “learning” in the cognitive radios, the challenge of survivability will not be trivial. Dense deployment of independent DSA networks, managed by autonomous and uncoordinated wireless service providers (WSPs) will have to fight valiantly for limited resources (spectrum), survival, efficiency, and quality of service (QoS). Moreover, although the FCC-proposed “commons” or “open access” paradigm that “allows unlimited number of unlicensed users to share frequencies”, the paradigm “does not provide any right to protection from interference” [2]. Limited and dynamically available resources and “no right to protection from interference” in the open access DSA model bring forth a serious challenge of self-coexistence among the secondary networks [10–14]. To make matters worse, there are risks of disruption of protocol-compliant CR devices/networks from malicious CR devices/networks. The “open” philosophy of the cognitive radio paradigm makes them more susceptible to various non-traditional spectrum etiquette attacks [12,15,16]. There are several ways in which operations of cognitive radio networks can be jeopardized. Cognitive disruption through malicious channel fragmentation/aggregation/bonding, or spectrum stealing through induced attack are just some of the unique mechanisms in DSA that can severely cripple the cognitive radio networks.

In this research, we particularly focus on a special vulnerability in cognitive radio networks with heterogeneous spectrum bands, known as induced attack. It is a special case of disruptive attack where an attacker uses its intelligence to force secondary users (or networks of users) to leave bands in the CR network. Cognitive radio networks are envisioned to be learning from their environment by observing, taking feedback, and analyzing their benefits from the band(s) through various course of actions. As such, inducing them maliciously through “intelligent” shadow-disruptive attacks has serious and long-term effects. For instance, successful induced attacks on candidate channel(s) will not only force the secondary networks to leave the band(s), but such successful frequent attacks may also harm their capability to learn about the particular bands. Therefore, even though, in reality, the spectrum bands may be available or even have high payoffs, secondaries will be reluctant to consider these bands as potential candidate channels, thus limiting their available radio resources. In the presence of such adversarial scenarios, the problems of survivability, network management, and specifying performance bounds become highly challenging that must be addressed; otherwise, the performance of the secondary networks will be degraded, defeating the purpose of the DSA paradigm.

In this paper, we formulate the problem as a game between SUs and the induced attacker, where the game is played in a heterogeneous dynamic spectrum access environment. The induced attacker’s aim is to not only disrupt the CRN operation but also to move the CRN toward “inappropriate cognition under malicious stimuli”. To investigate the conflict, we first demonstrate a static scenario where the usage of channels with different utilities is formulated as a solution of a zero-sum game. We further enhance our formulation for more realistic scenarios, when the channel utilities are no longer stationary and change abruptly (regime change) due to different channel characteristics, primary user, and disruptions that are not known beforehand. Under such dynamic regime changes, it becomes particularly complex to estimate and decide optimally amidst the presence of induced attacker. To address the difficulty, we adopt a tradeoff strategy between “exploration” and “exploitation” and compare our mechanism to benchmark situation where utilities are known exactly as soon as a regime change is detected.

The main contributions of this paper are as follows:

- A game-theoretic framework for making choices over channels to maximize channel utility in the presence of malicious induced attacks, this includes a closed form solution for optimal CRN strategies for both SU and attacker. We will often refer to SU as simply user.
- A stochastic model for defining, estimating, and learning channel utilities.
- Demonstration by analysis and simulation that the above obtained CRN strategies combined with estimation and learning can still provide adequate performance when actual utilities are replaced by their estimated values. In addition, we compare our mechanism to other conceivable benchmark strategies and show improved performance.

The rest of this paper is organized as follows. In Section 2, we review the related literature. Section 3 describes the IEEE 802.22 based system model, its challenges and the problem statement as the long term maximization of profit given channel utilities (leading to the game theoretic approach). In Section 4, we present the game theoretic approach. In Section 5, we describe a possible notion of utility. We further develop our study in Section 6, where utilities are no longer stationary for various reasons, and also not known beforehand. We present our proposed mechanism combining estimation and learning mechanisms with optimal play (from Section 4). Section 7 presents the simulation and numerical results. Conclusions are drawn in the last section.
2. Related works

While other aspects of CR networks have received significant attention over the past decade, understanding and addressing the security issues have remained particularly complex and challenging. The threats of using malicious cognitive radio(s) in DSA networks are even more prevalent and dangerous for several reasons: (a) They are highly “mobile” in every possible aspect due to the characteristics of software reconfigurability; (b) CR/DSA networks are susceptible to attacks ranging from passive eavesdropping to active interfering and frequent break-ins due to their open, ubiquitous, and interoperable nature; (c) Due to the open source nature of CR/DSA networks, it is practically impossible to establish a standard database to record the identity information of every CR node. Research in the area of CR network security is emerging in the recent times with some outstanding research contributions.

The security vulnerabilities in IEEE 802.22, including its security sub-layer, and the effects of various DoS attacks on the performance of 802.22 networks were briefly discussed in [18,19]. Clancy et al. [18] and Bhattacharjee et al. [12] also described a new class of attacks in CR networks, by which the secondary users can be trained to respond inappropriately to several stimuli. They give some specific examples of these attacks although no mitigation approaches are provided. In addition, more general discussions about the security issues in CR networks are given in [4,7,20–23]. However, most of the work mentioned above falls under review articles rather than a thorough and comprehensive analysis of defense mechanisms against those security issues.

Recently, several research groups have investigated specific DoS attacks in CR Networks. Chen et al. [24–27] described the Byzantine failure problem in the context of data fusion in the cooperative spectrum sensing. In this Byzantine attack, a malicious node intentionally sends falsified local spectrum sensing reports to the data collector in an attempt to cause the data collector to make incorrect spectrum sensing decisions. A novel reputation based mechanism called Weighted Probability Ratio Test was proposed to improve the robustness of data fusion against attacks. Zhang et al. [28] have provided a thorough survey on the defense mechanisms against Byzantine attacks. Another popular attack, drawing much attention, is the primary user emulation (PUE) attack [20,29], which was originally introduced by Chen et al. [30]. In the PUE attack, one or multiple attacking nodes transmit in forbidden time slots and effectively emulate the primary user to make the protocol compliant secondary users erroneously conclude that the primary user is present and evacuate that spectrum band. In order to thwart this attack, a localization based defense method was developed in [31], in which a non-interactive localization scheme is employed to detect and pinpoint the PUE attack. However, these localization mechanisms require a dedicated sensor network which may not be available in practical distributed DSA networks. Jin et al. [32] proposed a hypothesis based approach to mitigate PUE attacks using an analytical model for the received power at the secondary users, without assuming any prior knowledge about the position of either the malicious or the secondary users. The first analytical model for the received power was proposed in [33].

Radio Jamming is another common and disruptive DoS attack in wireless networks. In [34–37], Sampath et al. showed that jamming attackers can utilize CRs’ fast channel switching capability to amplify their jamming impact across multiple channels using a single radio. Later, a security-enhanced virtual channel rendezvous algorithm was proposed in [38] to improve the robustness of a DSA network against smart jamming attacks. Ma et al. discussed the jamming and anti-jamming procedures in multichannel CR systems [39]. Defense mechanisms against jamming based DoS attacks include Channel Surfing, Creating beamnull toward a jammer, Spatial Retreat, Mapping Jammed Region, Spread Spectrum, Honeynet based decoy, etc. [23,34,35,40–49]. Lee et al. discussed the detection of location spoofing attacks [50]. Dong et al. studied the impact of attacks on a network induced by wormholes [51]. A brief survey on defense mechanism for securing physical layer communications for CRN can also be found in [46].

Despite all of the above, there is still no framework that addresses one of the biggest vulnerability in cognitive radio networks: induced-attack [4,12,52,53]. Induced attack was envisioned in [15] and the severity of induced attack was in other survey papers [12]. However, there has not been any research paper exploring the vulnerability of this attack or a defense mechanism against it. This work is an attempt to analyze and devise stochastic defense mechanisms against induced attack in dynamic spectrum access environment.

3. Preliminaries and problem setting

3.1. System model

The IEEE 802.22 standard committee has aimed to develop the standard for the cognitive radio access strategy [54,55]. The standard specifies the physical and MAC layer operation of SUs in TV broadcast bands. A typical CR network is a single hop point-to-multipoint wireless networks, in which a central controller controls the resource allocation. An example of a CR network is shown in Fig. 1. Commonly, a base station (BS) determines which spectrum to use after all the customer premises equipment (CPE) under its supervision send the spectrum sensing report. The standard supports dynamic spectrum access where the SUs apply cognitive capability and use spectrum in an opportunistic manner. Both the BS and CPE perform spectrum sensing periodically to sense the presence of PUs. In Fig. 2, we depict the time domain representation of sensing and transmission period. The spectrum sensing reports are fused together to obtain the spectrum occupancy and availability map for the entire cell. Although the specifications provide strict protection mechanism to keep the incumbent primary
users free of interference, it does not confirm suitable protection mechanism for a CR network from another CR system. When multiple unlicensed operators are operating over a small available band of frequency, there is a chance that they will cause interference among themselves [4].

When a CR node switches on or moves to a new frequency channel, it performs listen before talk to detect the presence of the PU as well as BSs within its communication range. Since it is possible for each node in the network to choose its spectrum band, it is necessary for the given CR node to listen to the preferred channels of the BSs. The different physical propagation characteristics of electromagnetic waves over different spectrum bands is another concern for CRNs. A low-frequency signal (e.g., 700 MHz) can travel farther, penetrate walls and other obstacles but its information capacity is lower, and the accuracy in determining the direction of arrival is poorer. However, a higher frequency signal (e.g., 5.0 GHz) can only travel a shorter distance, but will be able to carry more information and will exhibit better directionality. Thus, the channels provide different reward or utility upon its usage.

3.2. Threat model

The "open" philosophy of the CR paradigm makes such networks susceptible to attacks by smart malicious users that could even render the legitimate CR spectrum-less. Due to software reconfigurability, CRs can even be manipulated to disrupt other CRNs or legacy wireless networks with greater impact than traditional hardware radios [7].

In the self-coexistence battle for spectrum opportunities, when the secondary networks are already competing for their survival against other secondaries, the effect of malicious disruptions can be even more fatal as there is no way to understand whether the disruptions are unintentional or intentional. The motivation for such shadow-disruptive attack behavior can be either monopolism, to capture as much spectrum as possible for themselves without maintaining any spectrum sharing etiquette and make other secondaries starve and eventually go out of the competition; or adversarial to disrupt other secondaries’ communications and shut them down (particularly applicable in environments filled with adversarial users/networks). To defend against such smart disruptions, it is absolutely critical to understand the uniqueness of the attack models/strategies in exploiting the finest granularity of spectrum agility and the shadow-disruptive nature of the malicious societies.

**Induced attack:** In this research, we investigate a particular type of disruptive attack where an attacker uses its intelligence to force users to leave bands in the CR network [12]. This is manifested by the awareness of the CR users that a potential attacker may exists, which could be reinforced by sensing disruptions (e.g. one could detect jamming attempts). As such, and due to the presence of these malicious disruptions, a user is often confronted to make a choice over channels (knowing that an attacker is likely to target the “best” ones). Therefore, to optimize some function, the user will have to migrate from using what is perceived as best channel to other channels. This not only affects performance, but also the ability of the user to effectively learn about channels. For example, few dynamic spectrum access algorithms gather channel access statistics for PUs in an attempt to predict when the channel will be idle [6]. Here the learning radios will be hindered by not being able to access their desired bands. Therefore, this degrades the performance of the CRN not only in terms of instantaneous transmission, but also on the long run.
For the remainder of this paper, the term “user” means a secondary user of the CRN. Similarly, the term “attacker” means a malicious secondary user of the CRN. For simplicity, the primary user of a channel is assumed to be “protected” and, hence, not affected by any attacks. The role of the primary user in our context is to simply claim or release the channel to be used by the secondary users (which include user and attacker).

Let us call \( u_j \) the utility of channel \( j \) when a user successfully transmits on that channel. Therefore, a successful transmission over channel \( j \) provides its user with a profit of \( u_j \). This utility \( u_j \) is a function of certain properties of the channel; for instance, it could be related to the transmission rate. However, if the channel is attacked during transmission, the utility is given by \( v_j \). Typically, \( u_j \) is positive and \( v_j \) is 0 or negative; however this is not necessary as long as \( v_j < u_j \). For instance, \( v_j \) could represent the energy loss (negative) in an unsuccessful transmission as a result of the attack. In such a case, the user makes a profit of \( v_j \) (instead of \( u_j \)). We assume that \( u_j \) and \( v_j \) have compatible units when they contribute to the profit. The goal of the user is then to maximize the long term profit by making an appropriate choice of channel when transmitting in each slot. The attacker has the opposite goal. The terms “user” and “attacker” are thus defined in this context. Consequently, an “attack” is regarded as any use of a channel by the attacker, i.e. with the intention to reduce the long term profit as described above.

While the model incorporates one user and one attacker, one could think of all users as one by the virtue that channel requests may go through a centralized controller [56,57]. Similarly, the aggregate behavior of all attacks may be thought of as coming from a single source. A user acting “optimally” in the face of such aggregate attacks does not need to explicitly distinguish the identity of the attacker. Nevertheless, the scenario of multiple users/attackers remains a valid one because “users” become “attackers” when competing for the same resources. This, however, is not the scope of this paper.

Our problem setting is focused on the following: We assume that the user is aware of the possible presence of an attacker, who in turn is regarded as a rational player. We consider the simple case where only one channel is chosen for transmission (by the user) and only one channel is chosen for an attack (by the attacker). The user and the attacker do not know which action their opponent will make at each time slot. With that in mind, we formulate the problem as a repeated zero-sum game, where both the user and the attacker make their choices over the available channels. The optimal strategy of our game depends on the knowledge of the channel utilities as described above. Realistically, however, such information is often not directly accessible. Thus we use a stochastic approach to estimate and learn the values needed. Both the game-theoretic approach and the stochastic estimation/learning are described in detail in the following sections.

4. The proposed game model

We formally study the case where only one channel is used at a time. We avoid the case when multiple channel are used at once to exclude the possibility that all channels can be attacked, which leads to a rather uninteresting scenario; we assume that no attacker has such a power [58–60]. We have \( n \) channels available in the CRN, a user, and an attacker. Channel \( i \) has two utilities \( u_i > v_i \) as explained above. We define \( \Delta_i = u_i - v_i > 0 \). In every time step, the user chooses a channel \( i \) for transmission, and a channel \( j \) is attacked. If \( i \neq j \), the user accumulates a profit \( u_i \); otherwise, the channel is blocked and the user accumulates a profit \( v_i \). As described in Section 3, the user seeks to maximize the total profit on the long run (while the attacker seeks the opposite).

It is not hard to see that pure strategies, where the choice over the channels is fixed, do not typically lead to Nash equilibrium. For instance, let \( i \) and \( j \) be the channels for the user and attacker, respectively. If \( i \neq j \), then the attacker may decrease the profit by following the user and redirecting the attack to channel \( i \). Similarly, if \( i = j \), then the user may increase the profit by moving away from that channel. This pattern of following and moving away can continue indefinitely from one channel to another, showing that an optimal strategy must be probabilistic in nature. In this case, the user/attacker generally seeks to maximize/minimize the expected profit.

In game theory, such a strategy, where each channel is chosen with some probability, is called a mixed strategy, and is the solution of a linear program for the zero-sum game with the following payoff matrix, where the entry in the \( i \)th row and \( j \)th column represents the profit when the user transmits on channel \( i \) and the attacker chooses channel \( j \).

\[
\begin{pmatrix}
  u_1 & u_1 & \cdots & u_1 \\
  u_2 & u_2 & \cdots & u_2 \\
  \vdots & \vdots & \ddots & \vdots \\
  u_n & u_n & \cdots & u_n \\
\end{pmatrix}
\]

We will analyze this game and obtain the optimal strategies for both user and attacker (in a Nash equilibrium sense). We will use \( p = (p_1, \ldots, p_n) \) and \( q = (q_1, \ldots, q_n) \) to denote the probabilities over the choice of channels for the user and the attacker, respectively. The goal is to compute \( p \) and \( q \) that represent the optimal mixed strategies: Given the optimal \( q \), any deviation from the optimal \( p \) will decrease the profit. Similarly, given the optimal \( p \), any deviation from the optimal \( q \) will increase the profit. Therefore, both user and attacker have the incentive to maintain these probabilities.

In principle, a knowledge of \( u \) and \( v \) means that the optimal strategies for the game are known to both user and attacker; therefore, it does not hurt if they actually announce them. On the one hand, if the user announces a strategy \( p \), the best response for the attacker will be to choose a channel \( j \) that achieves an expected profit of \( \min_j p_j v_j + \sum_{i \neq j} p_i u_i 
Please cite this article in press as: S. Mneimneh, et al., A game-theoretic and stochastic survivability mechanism against induced attacks in Cognitive Radio Networks, Pervasive and Mobile Computing (2017), http://dx.doi.org/10.1016/j.pmcj.2016.12.010
The user should choose $p$ to maximize the profit against that best response. Therefore, the user's interest is to maximize $\min_j p_jv_j + \sum_{i\neq j} p_iu_i$. This is equivalent to the linear program $LP_1$.

$$\begin{align*}
\text{maximize } & \quad z \\
\text{s.t. } & \quad \forall j, \quad z \leq p_jv_j + \sum_{i\neq j} p_iu_i \\
& \quad \sum_i p_i = 1 \\
& \quad \forall i, \quad p_i \geq 0.
\end{align*}$$

On the other hand, symmetrically, if the attacker has to announce a strategy $q = (q_1, \ldots, q_n)$, the best bet is to choose $q$ to minimize the expected profit under the user's best response, in other words, minimize $\max_i q_iv_i + \sum_{j\neq i} q_ju_i$. By observing that $q_iv_i + \sum_{j\neq i} q_ju_i = u_i - q_i\Delta_i$, we obtain the linear program $LP_2$ (the dual of $LP_1$):

$$\begin{align*}
\text{minimize } & \quad y \\
\text{s.t. } & \quad \forall i, \quad y \geq u_i - q_i\Delta_i \\
& \quad \sum_j q_j = 1 \\
& \quad \forall j, \quad q_j \geq 0.
\end{align*}$$

Let $z^*$ and $y^*$ be the optimal solutions for $LP_1$ and $LP_2$, respectively. It is known that $z^* = y^*$ by linear programming duality. Therefore, by solving $LP_1$, the user (maximizer) can determine a strategy for itself that guarantees an expected profit of at least $z^*$, no matter what the attacker does. And by solving the dual $LP_2$, the attacker (minimizer) can guarantee that the expected profit is at most $y^*$, no matter what the user does. Since $z^* = y^*$, this determines the optimal way to play for both.

What do we expect from the optimal solutions? Intuitively, some channels should never be used. For instance, given two channels $i$ and $j$ with $v_i = 0$, $u_i = 1$, $v_j = 2$, and $u_j = 3$, one would expect that the user should never choose channel $i$, as channel $j$ offers a better profit, even when under attack. Consequently, the attacker should also have no interest in attacking channel $i$. On the other hand, a user should typically favor a reliable channel $i$ with a small $\Delta_i = u_i - v_i$, because the attacker cannot dramatically deteriorate its profit. Algorithm 1 illustrates the optimal solutions for $p$ and $q$ and reflects these observations. In an initial phase (lines 4–13), some channels are eliminated. Among the remaining channels, the user assigns in a second phase (lines 14–17) a probability to a channel $i$ that is inversely proportional to $\Delta_i$. We leave the detail of how the solutions are derived to the Appendix.

**Algorithm 1:** Optimal solutions for $LP_1$ and $LP_2$

1. $S \leftarrow [n]$
2. $p \leftarrow 0$
3. $q \leftarrow 0$
4. repeat
5. \hspace{1em} $T \leftarrow \emptyset$
6. \hspace{1em} for $j \in S$ do
7. \hspace{2em} $x_j \leftarrow 1 - \frac{|S| - 1 + \sum_{i \in S} (v_i - u_j)\Delta_i^{-1}}{\Delta_j \sum_{i \in S} \Delta_i^{-1}}$
8. \hspace{2em} if $x_j < 0$ then
9. \hspace{3em} $T \leftarrow T \cup \{j\}$
10. \hspace{2em} end
11. \hspace{1em} end
12. $S \leftarrow S - T$
13. until $T = \emptyset$
14. for $j \in S$ do
15. \hspace{1em} $q_j \leftarrow x_j$
16. \hspace{1em} $p_j \leftarrow \frac{\Delta_j^{-1}}{\sum_{i \in S} \Delta_i}$
17. end

5. A notion of utility

Regardless of how channel utility is obtained, it is realistic to assume that such information is not readily available. In this section, we develop a stochastic model for utility. We assume here a standard paradigm of sensing and transmission,
though the detail is irrelevant for the theoretical treatment herein. In principle, a user can sense the channel to obtain some properties of the channel, such as availability and transmission rates. Consequently, the channel utility as perceived is related to a particular characteristic of the transmission over that channel. During the thth transmission period, the user observes an instantaneous profit $r_j(t)$ for the chosen channel $j$, when transmission is successful.

Let $I_j(t)$ be the indicator of the event that the user successfully transmits on channel $j$ during slot $t$. In other words, if $c(t)$ is the channel chosen by the user at time $t$, and $a(t)$ the one attacked, then $I_j(t) = 1_{(c(t)=a(t)=j)}$. Similarly, let $A_j(t) = 1_{(c(t)=a(t)\neq j)}$. Define $n_j(t) = \sum_{m=1}^{t} I_j(m)$. In the stationary operation of the system, the choice of channel is independent of the time slot, and the utility $u_j$ can be defined as a (stationary) expectation conditioned on a successful transmission:

$$u_j = \mathbb{E}[r_j(m) \mid I_j(m) = 1].$$

The statement below is a standard consequence of the law of large numbers, but we present it for completeness and to introduce a method for estimating $u_j$.

**Lemma 1.** Under stationary operation of the model, the utility satisfies

$$u_j = \frac{\mathbb{E} \left[ \sum_{m=1}^{t} r_j(m) I_j(m) \right]}{\mathbb{E}[n_j(t)]} = \lim_{t \to \infty} \frac{\sum_{m=1}^{t} r_j(m) I_j(m)}{n_j(t)}, \text{ w.p.1.} \quad (3)$$

The proof is left to the Appendix. Therefore,

$$u_j = \frac{\mathbb{E} \left[ \sum_{m=1}^{t} r_j(m) I_j(m) \right]}{\mathbb{E}[n_j(t)]}$$

is the quantity that we wish to estimate and, furthermore, it is worth noticing that for a finite $t$,

$$\mathbb{E} \left[ \frac{\sum_{m=1}^{t} r_j(m) I_j(m)}{n_j(t)} \right] \neq u_j.$$

The expression $\hat{u}_j = \sum_{m=1}^{t} r_j(m) I_j(m)/n_j(t)$ is the one usually used to estimate $u_j$ (at infinity they are equal with probability 1), thus we have a biased estimator because $\mathbb{E}[\hat{u}_j] \neq u_j$. If $\mathbb{P}(I_j(m) = 1)$ is known, then an unbiased estimation can be obtained by replacing the (random) denominator $n_j(t)$ by its expectation $\mathbb{E}[I_j(m)] = 1$. Typically, however, such information is hard to obtain in a dynamic environment, so we stick to the biased estimator (see Section 6).

Under induced attacks, if a chosen channel $j$ is attacked, then a collision happens and the transmission fails, and for simplicity, we now assume that a deterministic loss $e_j \geq 0$ is incurred due to the unsuccessful utilization of the channel. In this case, we set $u_j = -e_j$.

While $u_j$ and $u_j'$ (or more precisely their estimates) will determine the decisions of the players (as described in Section 4), the actual measured instantaneous profit of the user at a given time $t$ is given by $r_j(t) I_j(m) - e_j(t) A_j(m)$, and the average profit up to time $t$ is

$$G = \frac{1}{t} \sum_{m=1}^{t} \sum_{j=1}^{n} [r_j(m) I_j(m) - e_j(m) A_j(m)]. \quad (4)$$

In the following section, we look at the case where the utilities $u_j$ are estimated dynamically from the instantaneous values of $r_j(t)$ with noise, and we incorporate into the model the possibility that these values may undergo abrupt changes, as is the case for example when a primary user PU claims or liberates a channel.

### 6. Dynamic stochastic model

In the previous section, the game model assumes that the profit of accessing a channel is constant and known to both the user and the attacker. However, in reality, the utilities may vary over time due to noise and other considerations, such as the PU’s utilization of the channel. Nevertheless, a player who does not know the actual utility is typically able to estimate it after accessing the channel. In the following section, we describe how the user and the attacker can estimate the channel utilities in a realistic manner without any prior knowledge.
6.1. Statistical learning for dynamic estimation of utilities

If the expected utilities $u_j$ and $v_j$ were known exactly, then both user and attacker would be best using the optimal strategies found in Section 4. We assume here without loss of generality that $v_j$ is constant for all channels ($v_j = 0$ is our simulation) and we now use the notation $p(u)$ and $q(u)$ for the optimal strategies to make it explicit that they depend on the values of the expected utilities $u$.

At each time slot, the secondary users first sense the number of available channels $n$ (although this number depends on $t$, we use $n$ instead of $n(t)$ for ease of notation). Consider a stationary random strategy of the form

$$
\theta_j = \mathbb{P}(c(t) = j); \quad j = 1, \ldots, n
$$

$$
\alpha_j = \mathbb{P}(a(t) = j); \quad j = 1, \ldots, n
$$

where $c(t)$ and $a(t)$ are as defined in Section 5. During time slot $m$, when $I_j(m) = 1$ (also defined in Section 5), the user will measure the instantaneous profit

$$
r_j(m) = u_j + \eta_j(m)
$$

where $\{\eta_j(m) : m \geq 1\}$ are i.i.d. zero-mean random variables with bounded variance, representing the noise in actual observations of the channels’ properties. Given a stationary regime, the user has an estimate of $u_j$ at time $t$:

$$
\hat{U}_j(t) = \frac{1}{\theta_j(1 - \alpha_j)t} \sum_{m=1}^{t} r_j(m) I_j(m).
$$

This would be an ideal unbiased estimator because $E[\hat{U}_j] = u_j$ by Lemma 1 and the fact that $E[\eta_j(t)] = \theta_j (1 - \alpha_j) t$. However, if the user does not know $\alpha_j$ (because the attacker’s estimate of the utilities is unknown to the user), then the user will estimate $u_j$ at time $t$ as the sample average (as described in Section 5):

$$
\hat{u}_j(t) = \frac{1}{\eta_j(t)} \sum_{m=1}^{t} r_j(m) I_j(m).
$$

Although this latter estimator (7) is biased, it converges with probability one to $u_j$ under a stationary regime as $t \to \infty$ (that is, it is a consistent estimator). The reason for the bias lies in the $\mathbb{E}[1/\eta_j(t)] \neq 1/E[\eta_j(t)]$.

Similarly, the attacker defines its own analogous (unbiased and biased) estimators $\hat{U}_j(t)$ and $\hat{u}_j(t)$ by exchanging the roles of $\alpha_j$ and $\theta_j$, and $c(t)$ and $a(t)$.

Under a stationary regime, although different, both $\hat{U}_j(t)$ and $\hat{u}_j(t)$ are consistent estimators of $u_j$ for every $j$ that satisfies $\theta_j \neq 0$ and $\alpha_j \neq 0$ respectively (because user and attacker actually use the channel and, hence, obtain some estimates).

In the real system, however, utilities are not stationary and may occasionally change abruptly due to changes in channel characteristics coming from the PUs. We will assume that detection of channel availability/unavailability is instantaneous: as soon as the user and attacker sense the PU, they reset the current list of available channels. Similarly, when a PU liberates a channel the user and attacker have instantaneous knowledge. However, users have no means to know when the utilities themselves change values abruptly (e.g. when PU liberates a channel), except by estimation. This calls for a model for regime changes [61], where monitoring the changes becomes necessary in order to produce accurate statistics. In [62], for instance, we declare a regime change on channel $j$ whenever $w$ consecutive observations fall outside the region $\bar{u}_j(t) \pm 3 \sqrt{\text{Var}(u_j)}$, where $\text{Var}(u_j)$ is the estimated variance of $u_j$.

Once a regime change is declared, estimation of channel utilities and their variances is reset for all channels. Then choosing the channel more often will result in a faster and more accurate estimation of the channel’s utility, and thus also of $p(u)$ and $q(u)$. However, the real goal is to play the game optimally, and not to estimate optimally. That is, we wish to estimate $p(u)$ and $q(u)$ accurately, so that the actions are chosen precisely by setting $\theta = p(u)$ and $\alpha = q(u)$. Unfortunately, using estimates in real time does not ensure convergence, since $p(\hat{u}_j(t))$ or $q(\hat{u}_j(t)$ might be zero for some $j$. Specifically, using

$$
\theta_j(t) = p(\hat{u}_j(t))
$$

$$
\alpha_j(t) = q(\hat{u}_j(t))
$$

may lead to a very bad strategy, and it will not converge to the real values if at the moment of a regime change the channel in question satisfies $p(\hat{u}_j(t)) = 0$. In this case, channel $j$ will not be chosen for transmission and, consequently, the estimation will never be updated. We adopt a trade-off strategy between exploration and exploitation, as described below:

As soon as a regime change is detected for any channel, the user/attacker declares a learning period of length $T$ time slots (exploration). At this point, the algorithm resets $t = 0$ and uses uniform probabilities for sampling the channels. After the training period is over, we revert to the probabilities given by the newly available estimates $p(\hat{u}_j(t))$ and $q(\hat{u}_j(t))$ (exploitation).

For simplicity, our simulations limit regime changes to changes in PU activity. The following section provides simulation results to evaluate our game-theoretic solution and stochastic estimation mechanism against other reasonable benchmark strategies.
Algorithm 2: Learning algorithm for user

1. learning ← Tc
2. while True do
3.   if change in PU activity then
4.     learning ← Tc
5.     t ← 0
6. end
7.   t ← t + 1
8.   if learning > 0 then
9.     access available channels with equal probability
10.    update ūi as in (7)
11.    learning ← learning − 1
12. end
13. else
14.    access channel i with probability pi from Alg. 1
15.    update ūi as in (7)
16. end
17. end

7. Simulation results and discussions

For the simulator, a sensing period followed by a transmission period form a slot. The user and the attacker are synchronized for the sensing and transmission periods.

The total number of channels is fixed. For each channel, the activity of its PU is modeled as an on–off process (each experiment will describe the parameters for that process).

The utility for a channel (ui) is 0 when its PU is present, as channel j is not accessible within the CRN by another secondary user. Both the user and the attacker can sense the spectrum during the sensing period of each slot and generate a list of available channels to access. They choose one channel to access at the starting of the transmission period. That choice remains persistent during the entire time slot.

If the channel j is accessed at time slot m without any collision, then a profit rj(m) as described in (5) is observed. This is true for both user and attacker. For simplicity, the simulator uses ηj as a uniform random variable \( U(−1, 1) \) for all channels and utility values. That is, the sequence ηj(m) consists of independent and identically distributed zero-mean uniform random variables. Both the user and the attacker observe \( v_j = 0 \) when they collide. They both keep track of the channels used and their observed profits in order to estimate \( û_j(m) \) and \( ū_j(m) \) in accordance with (7).

Fig. 3 provides a snapshot of a typical simulation, where the x-axis represents time slots. The PU process was as follows: The number of time slots that a PU stays active is a uniform random variable on [50, 150]. The number of time slots without PU activity is independent of previous activity and is uniform on [50, 600]. The colors blue, green, red, and cyan correspond to channels 0, 1, 2 and 3 respectively.

The first subplot sketches the actual utility \( u_j(m) \) of channel j during slot m. In this subplot, we can see abrupt changes in the utility of channels due to PU arrival or departure. When a channel is being used by the PU, its utility is 0 as described earlier.

The second subplot shows the channels accessed by the user and the attacker in each time slot. Blue circles indicate the channels used by the user, and red squares trace the channels used by the attacker.

The third subplot reports the observed profit \( r_j(m) \) for the user. If at time slot m, the user and the attacker access the same channel j, the user observes \( v_j(m) = 0 \); otherwise, \( r_j(m) = u_j(m) + η_j(m) \); where \( η_j(m) \) is the noise.

The fourth subplot illustrates the estimated utility \( û_j(m) \) of the user for the different channels during the simulation. A yellow shaded region indicates a learning period. It can be observed that the estimated utilities change abruptly during learning periods, and slowly otherwise.

The fifth subplot portrays the probabilities \( θ_i(m) \) with which the user accesses channels at the given time slot. During a learning period, \( θ_i(m) \) is uniform for all channels; otherwise, the user chooses channels using \( θ_i(m) = p(û_i(m)) \) during the exploitation phase. The effect of induced attack in this subplot can be clearly observed. Although channel 3 has the highest utility, the attacker forces the user to choose channel 3 with lower probability as dictated by the strategy.

The sixth subplot depicts the channel probabilities \( α_j(m) \) for the attacker, in a way similar to the fifth subplot.

In this simulation, we have used a synchronized attacker, i.e. the learning period for both the user and the attacker start and end simultaneously (here \( T_c = T_a = 30 \)). Our basic assumption is that both the user and the attacker can sense PU arrival and departure at the same time. This assumption makes them start and end the learning period in a synchronized way.

We outline below several experiments based on the general simulation framework thus described.
7.1. Experiment 1: learning times

Under the presence of noise, both the user and the attacker enter a learning period to estimate channel utilities as soon as they detect changes in PU activity. In this experiment, we investigate how the length of the learning period affects the overall performance. In particular, we explore the possibility of an equilibrium point \((T_c, T_a)\), from which both user and attacker do not wish to deviate.

The simulator uses 4 channels with utilities 6, 18, 28 and 40. The number of time slots that a PU stays active is a uniform random variable on \([25, 100]\). The number of time slots without PU activity is independent of previous activity and is uniform on \([100, 400]\).
For each combination of user–attacker parameters \((T_c, T_a)\), we run 25 simulations of 100,000 time slots each to obtain averages. Fig. 4 shows the average profit of the user with different learning periods in a 3-dimensional plot as a function of \(T_c\) and \(T_a\). The average profit is calculated as in (4) for \(t = 100,000\).

As before, the probability of a channel is uniform in a learning period, and is obtained by \(p(\hat{u})\) and \(q(\tilde{u})\) for the user and attacker respectively during exploitation. The probabilities computed by one player are unknown to the opponent.

On the one hand, if the user has long learning periods, it accesses all channels uniformly for long times, which does not allow the user to effectively explore the better channels. On the other hand, short learning periods do not enable the user to learn the channel utilities properly.

The same observation holds in case of the attacker. Longer learning periods do not allow the attacker to effectively jam the better channels, while shorter learning periods do not enable the proper learning of the channel utilities.

In deed, we observe that there is a saddle point \((T_c, T_a)\), as indicated in Fig. 4; any deviation from this point hurts the player in question; a deviation in \(T_c\) decreases the average profit, and a deviation in \(T_a\) increases it. Fig. 4 reveals the optimal learning period of 6 time slots for both the user and the attacker given the particular simulation parameters. In general, they need not be the same. An approach for computing \((T_c, T_a)\) is beyond the scope of our paper. Practically, the learning time may be adjusted in an ad-hoc way until a better performance is observed. The following section illustrates the change in performance as a function of \(T_c\).

7.2. Experiment 2: a stronger adversarial model

Following up on the previous experiment, we consider a stronger adversarial model in which the attacker has accurate information of the utilities \(u_j\), and thus does not require a learning mechanism to estimate them. We call this here the game with accurate information, in contrast to estimated information as before. For this simulation, we fix \(T_a = 6\) as obtained above and vary \(T_c\). The plot in Fig. 5 of the game with estimated information is consistent with the previous results, showing a peak for the average profit when \(T_c = 6\). The game with accurate information is not a practicality, even when the accurate information is on the attacker side, but it represents a worst-case scenario for the user and provides a conservative view of performance.

Naturally, the length of the optimal learning period varies with the number of channels, the frequency of PU arrival/departure, and the mean and variance of channel utilities (given by \(u_i\) and \(\eta\)).

7.3. Experiment 3: benchmark strategies

To evaluate the efficiency of the game model, we compare it with other typical channel selection schemes for a number of channels \(n \in [2, 20]\). The channel utilities are generated for each channel independently and uniformly at random in \([6, 30]\). The number of time slots that a PU stays active is a uniform random variable on \([75, 125]\). The number of time slots without PU activity is independent of previous activity and is uniform on \([300, 500]\).

In the **Random** channel selection, both the user and the attacker choose channels uniformly at random (with equal probability). In the **Greedy** channel selection, the probability of selecting a channel is higher for the channels with higher utility. In this scheme, if \(n\) channels are available, they are ranked by increasing utility from 1 to \(n\). The probability of selecting channel \(j\) is \(j/\sum_{i=1}^{n} i\).
Fig. 6. Comparison of average profit to Random and Greedy benchmarks.

The Random selection scheme does not require knowledge of \( u_j \); therefore, to present a fair comparison among the different schemes, we may assume that both the user and the attacker have accurate information of channel utilities, i.e. no learning periods are required. We simulate the Random and the Greedy schemes with accurate information, and our game model with both accurate and estimated information (Fig. 6).

Fig. 6 shows the simulation results for the four schemes. With the Random scheme, the user does not utilize the best channels frequently enough; even a random attacker is sufficient to expose this deficiency. Greedy performs better, but the greedy attacker can still frequently prevent the user from successfully utilizing the best channels. Our game model with accurate information based on computing \( p \) and \( q \) performs the best.

Finally, our game model with estimated information is indeed the scenario of interest, where both user and attacker use learning to estimate utilities. The optimal learning periods \( T_c \) and \( T_a \) for the different number of channels were computed by separate simulations. Fig. 4 shows that our game model with estimated information falls between its accurate information counterpart and Greedy.

8. Conclusion

In this paper, we used a game to model the actions of choosing channels for transmission (by the user) and for attack (by the attacker). We described a closed form solution for the game when the channel utilities are known and fixed.

The generalization to non-stationary channel utilities is necessary in order to realistically model the PU activity and the uncertainty in obtaining the values of those utilities. In our formulation, we assume that channel utilities can change abruptly, not only by the activity of the PU, but also due to other factors such as channel deterioration, and we allow for the monitoring of regime changes (although for simplicity we only allow changes due to PU activity in our simulation).

Because the optimal game strategies depend on utilities, uncertainty in the estimation of the utilities may introduce significant bias in the operation of the game when the (noisy) estimates are used directly to compute the probabilities \( p \) and \( q \) for the user and the attacker, respectively. In particular, just after a regime change, channels may have bad estimates of their utilities. In this case, using the closed form solution as if the estimates were the true values may lead to very bad strategies. Instead, we explore a mechanism to learn the new values of the utilities before using the closed form solutions. This mechanism relies on accessing channels in a uniform way during a learning period \( T \). Our simulation results show that the game theoretic solution (exploitation phase) combined with the estimates of utilities through learning (exploration phase), lead to improved results when compared to some benchmark strategies.

Appendix

Proof of Lemma 1. Consider the following equality:

\[
\mathbb{E}[r_j(m) I_j(m)] = \mathbb{E}[r_j(m) \times 1 \mid I_j(m) = 1] P(I_j(m) = 1) + \mathbb{E}[0 \mid I_j(m) = 0] P(I_j(m) = 0).
\]
Therefore, the channel’s utility is defined by

\[ u_j = \mathbb{E}[r_j(m) \mid I_j(m) = 1] = \frac{\mathbb{E}[r_j(m)I_j(m)]}{\mathbb{P}(I_j(m) = 1)}. \]

Under stationary operation, and from the definition of \( n_j(t) \), it follows that (linearity of expectation)

\[ t \mathbb{P}(I_j(m) = 1) = \mathbb{E}[n_j(t)]. \]

Therefore,

\[
\begin{align*}
\frac{\mathbb{E}[r_j(m)I_j(m)]}{\mathbb{E}[n_j(t)]} = & \frac{\mathbb{E}\left[\sum_{m=1}^t r_j(m)I_j(m)\right]}{\mathbb{E}[n_j(t)]} = \frac{\mathbb{E}\left[\sum_{m=1}^t r_j(m)I_j(m)\right]}{t},
\end{align*}
\]

where \( \mathbb{E}[r_j(m)I_j(m)] = \mathbb{E}\left[\sum_{m=1}^t r_j(m)I_j(m)\right]/t \) is given by the linearity of expectation.

The final step uses the strong law of large numbers for stationary processes to establish the almost sure convergence of both numerator and denominator:

\[
\begin{align*}
\mathbb{E}[r_j(m)I_j(m)] = & \lim_{t \to \infty} \frac{\sum_{m=1}^t r_j(m)I_j(m)}{t}, \\
\frac{\mathbb{E}[n_j(t)]}{t} = & \mathbb{E}[I_j(m)] = \lim_{t \to \infty} \frac{n_j(t)}{t}.
\end{align*}
\]

Because both numerator and denominator converge with probability one, the ratio also converges with probability one, which establishes the result.

**Optimal solutions for LP1 and LP2**

Given a set of channels \( T \subseteq [n] \), we define the linear program \( LP_T^1 \) to be the same as \( LP_1 \) after dropping all channels \( k \in T \), i.e. dropping \( p_k \) and the \( k \)th constraint for all \( k \in T \). This corresponds to \( LP_1 \) when the \( k \)th row and the \( k \)th column for all \( k \in T \) are eliminated from the payoff matrix. We define \( LP_T^2 \) similarly.

**A feasible solution for LP1**

A feasible solution for \( LP_1 \) can be obtained by making

\[ p_1v_1 + \sum_{i \neq 1} p_iu_i = p_2v_2 + \sum_{i \neq 2} p_iu_i = \cdots = p_nv_n + \sum_{i \neq n} p_iu_i \]

which requires solving

\[ p_i\Delta_i = p_{i+1}\Delta_{i+1} \]

\[ \sum_i p_i = 1 \]

and yields the solution:

\[ p_i = \frac{\Delta_i^{-1}}{\sum_j \Delta_j^{-1}} \]

for a profit \( z = \sum_i u_i\Delta_i^{-1} - 1)/\sum_i \Delta_i^{-1} \). In addition, dropping any number of channels can only yield a feasible solution for \( LP_1 \).

**Lemma 2.** A solution for \( LP_T^1 \) augmented with \( p_k = 0 \) for all \( k \in T \) is feasible for \( LP_1 \).

**Proof.** The following linear program is \( LP_1^1 \):

\[
\begin{align*}
\text{maximize} & \quad z \\
\text{s.t.} & \quad \forall j \not\in T, \quad z \leq p_jv_j + \sum_{i \not\in T \cup \{j\}} p_iu_i \\
& \quad \sum_{i \not\in T} p_i = 1 \\
& \quad \forall i \not\in T, \quad p_i \geq 0.
\end{align*}
\]
For a given \( l \in T \), the dropped constraint in \( LP_1 \) is \( z \leq p_jv_j + \sum_{i \in S} p_iu_i \). When \( p_k = 0 \) for all \( k \in T \), this constraint becomes \( z \leq \sum_{j \in T} p_jv_j \). But since \( u_i > v_i (\Delta_i > 0) \), this constraint is dominated by any constraint in \( LP^T_1 \) that replaces its sum the term \( p_jv_j \) with the term \( p_j\bar{v}_j \) for some \( j \), as shown above. ■

A feasible solution for \( LP_2 \)

Consider a relaxed version of \( LP_2 \) by dropping the positive constraint on \( q \) (it has been replaced by \( x \) below to emphasize that the two are different solutions and for the purpose of maintaining a clear notation):

\[
\begin{align*}
\text{minimize} & \quad f \\
\text{s.t.} & \quad \forall i, \quad f \geq u_i - x_i \Delta_i \\
& \quad \sum_j x_j = 1
\end{align*}
\]

where \( q \in [0, \infty) \) has been replaced by \( x \in \mathbb{R} \).

**Lemma 3.** The optimal solution \( f^* \) for the above linear program can be obtained by making

\[
\begin{align*}
u_1 - x_1 \Delta_1 = u_2 - x_2 \Delta_2 = \cdots = u_n - x_n \Delta_n.
\end{align*}
\]

**Proof.** To show this, assume that \( \sum_j x_j = 1 \) and the above equality holds. A better solution will have to decrease \( u_i - x_i \Delta_i \) for every \( i \). Since \( \Delta_i > 0 \), this means \( x_i \) must increase for every \( i \), making it impossible to maintain \( \sum_j x_j = 1 \). ■

With the addition of \( \sum_j x_j = 1 \), the equality in Lemma 2 yields:

\[
x_j = 1 - \frac{(n - 1) + \sum_i (v_i - v_j) \Delta_i^{-1}}{\Delta_i \sum_i \Delta_i^{-1}}
\]

for \( f^* = [(n - 1) + \sum_i v_i \Delta_i^{-1}] / \sum_i \Delta_i^{-1} \). Observe that \( y^* \geq f^* \) (because \( x \) is less constrained than \( q \)), but if \( x \geq 0 \), then \( y^* = f^* \) and we have an optimal solution for \( LP_2 \).

**Lemma 4.** Let \( T = \{k|x_k < 0\} \). A solution to \( LP^T_2 \) augmented with \( q_k = 0 \) for all \( k \in T \) is feasible for \( LP_2 \).

**Proof.** The following linear program is \( LP^T_2 \).

\[
\begin{align*}
\text{minimize} & \quad y \\
\text{s.t.} & \quad \forall i \not\in T, \quad y \geq u_i - q_i \Delta_i \\
& \quad \sum_{j \not\in T} q_j = 1 \\
& \quad \forall j \not\in T, \quad q_j \geq 0.
\end{align*}
\]

Let \( l = \max_{i \in T} u_i \). Since \( x_k < 0 \) and \( \Delta_i > 0 \) and \( f \geq u_i - x_i \Delta_i \), it follows that \( f^* > u_i \); therefore we have \( y^* \geq f^* > u_i \). For a given \( k \in T \), the dropped constraint in \( LP_2 \) is \( y \geq u_k - q_k \Delta_k \). When \( q_k = 0 \), this constraint becomes \( y \geq u_k \), which is dominated by the same constraint when \( k = l \). Therefore, to prove the claim, we need to show that the dropped constraint \( y \geq u_i \) is dominated by another in \( LP^T_2 \), i.e. that \( u_i \geq u_k - q_k \Delta_k \) for some \( i \not\in T \). Assume \( u_i > u_k - q_k \Delta_k \) for all \( i \not\in T \), then we have a feasible solution \( y = u_i \) for \( LP_2 \), a contradiction since \( y^* > u_i \). ■

The analysis above suggests the following approach for finding a feasible solution for \( LP_2 \): Given the optimal solution to the relaxed version of \( LP_2 \) (4), drop all channels in \( T = \{k|x_k < 0\} \), thus obtaining a new instance \( LP^T_2 \) with a smaller number of channels. Do this repeatedly until no channel \( j \) satisfies \( x_j < 0 \). Finally, solve \( LP_2 \) by making \( q_j = x_j \) for every \( x_j \geq 0 \) and \( q_j = 0 \) for every dropped channel \( j \).

**The optimality of both solutions**

Drop from \( LP_1 \) exactly those channels that are dropped from \( LP_2 \), and let \( S \) be the set of channels that remain, observe that

\[
\begin{align*}
z &= \frac{\left( \sum_{i \in S} u_i \Delta_i^{-1} \right) - 1}{\sum_{i \in S} \Delta_i^{-1}} = \frac{(|S| - 1) + \sum_{i \in S} (u_i \Delta_i^{-1} - 1)}{\sum_{i \in S} \Delta_i^{-1}} \\
&= \frac{(|S| - 1) + \sum_{i \in S} v_i \Delta_i^{-1}}{\sum_{i \in S} \Delta_i^{-1}} = y.
\end{align*}
\]

Please cite this article in press as: S. Mneimneh, et al., A game-theoretic and stochastic survivability mechanism against induced attacks in Cognitive Radio Networks, Pervasive and Mobile Computing (2017), http://dx.doi.org/10.1016/j.pmcj.2016.12.010
Thus the feasible solutions for $LP_1$ and $LP_2$ become optimal by linear programming duality and, therefore, the above is also equal to $z^*$ and $y^*$. This essentially translates to the following theorem.

**Theorem 1.** The optimal solutions for $LP_1$ and $LP_2$ can be computed by Algorithm 1 with $u > v$, and $n$ as input, for a profit of $z^* = y^* = \left(\sum_{i \in S} u_i \Delta_i^{-1} \right) / \left(1 / \sum_{i \in S} \Delta_i^{-1} \right)$, where $S$ is the set of channels $i$ with $p_i > 0$.

**References**


Please cite this article in press as: S. Mneimneh, et al., A game-theoretic and stochastic survivability mechanism against induced attacks in Cognitive Radio Networks, Pervasive and Mobile Computing (2017), http://dx.doi.org/10.1016/j.pmcj.2016.12.010


