A FICTITIOUS PLAY-BASED GAME-THEORETICAL APPROACH TO ALLEVIATING JAMMING ATTACKS FOR COGNITIVE RADIOS

Kresimir Dabcevic, Alejandro Betancourt, Lucio Marcenaro, Carlo S. Regazzoni

DITEN, University of Genova
Via Opera Pia 11A 16145 Genoa - Italy
{kresimir.dabcevic, alejandro.betancourt, mlucio, carlo}@ginevra.dibe.unige.it

ABSTRACT

On-the-fly reconfigurability capabilities and learning perspectives of Cognitive Radios inherently bring a set of new security issues. One of them is intelligent radio frequency jamming, where adversary is able to deploy advanced jamming strategies to degrade performance of the communication system. In this paper, we observe the jamming/anti-jamming problem from a game-theoretical perspective. A game with incomplete information on opponent’s payoff and strategy is modelled as a Markov Decision Process (MDP). A variant of fictitious play learning algorithm is deployed to find optimal strategies in terms of combination of channel hopping and power alteration anti-jamming schemes.

Index Terms— jamming, anti-jamming, cognitive radio, game theory, fictitious play, markov models, channel surfing, power alteration

1. INTRODUCTION

Software Defined Radios (SDRs) and Cognitive Radios (CRs) [1] have over the last decade emerged as potential solutions to spectrum underutilization problem. However, the introduced reconfigurability potentials and unique cognitive characteristics are also bringing a set of new security risks and issues [2]. Among them, Primary User Emulation Attacks [3, 4], Byzantine Attacks [5, 6] and Intelligent Jamming Attacks [7, 8, 9] have received particular attention from the research community. Radio frequency (RF) jamming refers to intentional creation of interference at the target receiver with the aim of disrupting communication. RF jamming has found particular application in military domain, where various jamming and anti-jamming systems were studied [10, 11].

Game theory - a study of decision making under competition - has recently sparked interest as a tool for mathematical formalization of the Intelligent Jamming problems. Using game theory makes it possible to model and analyze interactions between the transmitters and the jammers in the system, as their overall goals are typically negatively correlated. Authors in [12] have formulated the problem as a zero-sum stochastic game, where channel hopping was considered as the anti-jamming scheme, and minimax-Q as the learning mechanism. The method was extended in [13], comparing the results of Q-learning with those of the policy iteration scheme. In [14] and [15], authors considered multi-carrier power allocation as an anti-jamming strategy, and also formulated the games as zero-sum.

In this work, we extend upon the aforementioned ideas and formulate a game which takes into account both channel hopping and power alteration as defense strategies. By introducing the hopping and transmission costs, as well as diverse reward factors for the transmitter and jammer side, the game is formulated as non-zero-sum. Finding equilibrium points in stochastic non-zero-sum games such as this one is a non-trivial task. Hence, we focus on simulation results for finding near-optimal strategies for a game with incomplete information on user payoffs and strategy distributions. A variant of fictitious play online learning algorithm [16] is proposed for updating the stochastic distributions of such strategies.

To the best of our knowledge, this is the first game-theoretical contribution which considers an increased action space created by combining channel hopping and power alteration schemes. Starting from a naive game where players take their decisions retroactively, the motivation for switching to a proactive game is shown. A stochastic decisioning policy is proposed as optimal policy for fictitious play learning algorithm.

The remainder of the paper is organized as follows: section 2 describes the system model. Game formulation, along with evolution from the naive deterministic game to the proactive stochastic game is presented in section 3. Simulation results for a game with 2 channels and 2 discrete values of transmission power are presented in section 4, whereas conclusions and the roadmap are given in section 5.

2. SYSTEM MODEL

Consider a transmitter-receiver pair that is trying to maintain continuous communication over one of the $n_f$ pre-assigned channels, and a jammer that is trying to disrupt the communication by creating interference. All of the nodes are assumed
to be equipped with SDR / CR technology, which allows them
to alter their transmission power and transmission frequency
on-the-fly. Transmitter and receiver have the exclusive spec-
trum rights to all of the considered channels, and are equipped
with the ability to tune to the same channel at a given time
instance using the pre-defined pseudo-random pattern. Jammer
is using narrowband waveforms for creating interference, al-
lowing it to create interference only at a single channel at a
time. Furthermore, it is equipped with spectrum sensing ca-

pabilities [17], which allows it to discover which channel is
currently being used by the transmitter, and consequently
to start creating interference at a given channel.

To mitigate effects of jamming, and increasing Signal
to Interference plus Noise Ratio (SINR) at the receiver to
the level needed for successful decoding, transmitter has the
choice of either changing its transmission frequency (chan-
nel hopping) [18], or transmitting at a higher power (power
alteration).

3. GAME FORMULATION

Analyzing RF jamming and anti-jamming strategies is a com-
plex problem that depends on multiple factors, some of which are
time-varying and channel-dependant. In order to approach
the jamming/anti-jamming problem from a game-theoretical
perspective, a set of assumptions and abstractions has to be
taken, such as: i) the available channels are non-overlapping
and perfectly orthogonal, i.e. jamming one channel has no
effect on the neighboring channels. ii) the available channels
are stationary and frequency-flat. iii) jamming is modelled as
a discrete event, i.e. it either occurs with success or failure
iv) all the players in the game maintain their relative posi-
tions as well as antenna orientations (radiation patterns) with
respect to each other.

Following these assumptions, a multi-stage game with
two players is modelled. At the end of each step, each player
receives his immediate payoff for the current step, and takes
a decision regarding his action in the following step. The
decision is two-dimensional, as the player needs to decide on
both his frequency and transmission power in the next step.

The modelled game takes into account reward for the suc-
cessful transmission or jamming, as well as the costs of fre-
cquency hopping and the transmission cost. Payoff at the end
of the step s for transmitter T is given as:

\[ P_s^T (C_s^T, f_s^T, f_f^T) = R_s^T \cdot (1 - \alpha) - \gamma \cdot C_s^T \]  \( \text{(1)} \)

Here, \( R_s^T \) is the reward for successful transmission, \( H \) is the
fixed cost of hopping, \( C_s^T \) is the transmitter’s current cost of
transmission, \( f_s^T \) is the frequency that transmitter is currently
using. \( \alpha \) denotes event of successful transmission (2), and
\( \beta = 1 \) if the transmitter decides to hop and \( \beta = 0 \) otherwise.

\[ \alpha = \begin{cases} 
1 & \text{if } C_s^T > C_f^T \text{ or } f_s^T \neq f_f^T \\
0 & \text{if } C_s^T \leq C_f^T \text{ and } f_s^T = f_f^T 
\end{cases} \]  \( \text{(2)} \)

Similarly, jammer J’s payoff at the step s is given as:

\[ P_s^J (C_s^J, f_s^J) = R_s^J \cdot (1 - \alpha) - H \cdot (1 - \gamma) \cdot C_s^J \]  \( \text{(3)} \)

\( R_s^J \) is jammer J’s reward for successful jamming, \( C_s^J \) jammer’s
cost of transmission in s, and \( \gamma = 1 \) if jammer decides to hop
and 0 if it does not.

In each step transmitter and jammer can deploy \( 1 \leq C_s^T \leq C_s^{MAX} \) and \( 0 \leq C_s^J \leq C_s^{MAX} \) with \( C_s^{MAX} \leq C_f^{MAX} \).

3.1. A naive deterministic game

First, a naive deterministic game is modelled. At the end of
each step s, each player observes the current payoff, transmis-
sion power and transmission frequency. In case that the play-
ers were transmitting at the same frequency, transmitter is also
able to estimate jammer’s transmission power - presumably
calculated from the SINR obtained at the receiver - whereas
jammer is always able to estimate transmitter’s power as well
as transmission frequency using the spectrum sensing mech-

anism. Then, given these observations, each player devises
an action that will maximize their payoff in the next state.
It is easy to show that the problem comes down to a simple
ternary decision. Each case denotes a simplified action set
for the transmitter (4) and jammer (5) as (power, frequency):

\[ A_{s+1}^T = \begin{cases} 
\text{(KS), if } \alpha(s) = 1 \\
\text{(RC), if } \alpha(s) = 0 \text{ and } (H < C_s^T + \Delta C^T \text{ or } C_s^T + \Delta C^T > T_s^{MAX}) \\
\text{(IS), if } \alpha(s) = 0 \text{ and } (H \geq C_s^T + \Delta C^T \text{ and } C_s^T + \Delta C^T \leq T_s^{MAX}) 
\end{cases} \]  \( \text{(4)} \)

\[ A_{s+1}^J = \begin{cases} 
\text{(KS), if } \alpha(s) = 1 \\
\text{(RC), if } \alpha(s) = 0 \text{ and } f_s^T \neq f_f^T \\
\text{(IS), if } \alpha(s) = 1 \text{ and } f_s^T = f_f^T 
\end{cases} \]  \( \text{(5)} \)

However, it is legitimate to expect that the learning mech-

anisms on either (or both) of the sides would allow the players
to take more advanced decisions, thus exploiting the decisions
of the opponent. Illustratory example of gradual evolution of
the game when such an arms race is present is shown in Fig. 1.
In time 1, both players are observing whether their action in the given step brought them positive payoff. If so, they choose the action (KS) for the following step, otherwise they keep increasing their transmission powers by \( \Delta C^T \) (transmitter) or \( \Delta C^J \) (jammer). This is repeated for as long as \( C^T_{s+1} < H \) and \( C^J_{s+1} \leq T_{\text{MAX}} \). State of the system is illustrated as \( T \) when transmission is successful and \( J \) when jamming is successful. Then, at time 2, transmitter decides to switch to another frequency. However, by observing jammer’s behaviour in time 1, it also realizes that better result would be yielded by proactively increasing its transmission power by two discrete increments in every step. When cost of transmission has once more risen above the cost of hopping, it will hop back to frequency 1 (or any other frequency). In time 3, jammer will observe this pattern and will decide to increase the probability of successful jamming by proactively increasing its transmission power in each step by 3 increments. Intuitively, the game will eventually evolve towards proactive hopping and transmitting with maximum power in every step.

3.2. A proposed game based on fictitious play

By observing history of the previously obtained payoffs for a given action and incorporating these observations into their decision-making process, players can obtain even better payoffs in the future. Fictitious play is an iterative algorithm where, at every step, each player takes the best response action that will optimize its payoff, given that other players take their actions independently at random according to the stochastic distribution of their own payoffs. Best response can pertain to choosing the action either deterministically or randomly with a certain stochastic distribution, depending on the adopted decisioning policy.

In each step \( s \), transmitter observes its current state \((C^T_s, f^T_s)\) and its possible actions \((C^T_{s+1}, f^T_{s+1})\), and compares the expected payoffs of each action by accessing a vector of expected payoffs \( P^T \). For \( n_T \) possible discrete values of transmission powers and \( n_f \) channels in the system, cardinality of the vector is \( |P^T| = (n_T \cdot n_f)^2 \). As transmitter and jammer are taking their actions simultaneously, the received payoff \( P^T_{s+1} \) depends on the jammer’s action in the current step. Once that transmitter receives its payoff, it updates \( P^T \) according to (6).

\[
P^T_{s+1}(C^T_{s+1}, f^T_{s+1}, f^J_{s+1}) = \frac{\sum_{i=1}^{n_T} \sum_{j=1}^{n_J} N(C^T_{s+1}, f^T_{s+1}, i, j) P(C^T_s, f^T_s, f^J_s, i, j)}{\sum_{i=1}^{n_T} \sum_{j=1}^{n_J} N(C^T_{s+1}, f^T_{s+1}, i, j)}, \tag{6}
\]

where \( N(C^T, f^T, C^J, f^J) \) denotes the number of times that the state \((C^T, f^T, C^J, f^J)\) has occurred during the game and \( P(C^T, f^T, C^J, f^J) \) is the payoff corresponding to that state.

Similarly, jammer updates its vector of expected payoff \( P^J \).

State transitions can be depicted by finite-state Markov chains, where transition probabilities change dynamically, depending on the available up-to-date history and the decisioning policy. Two decisioning policies (greedy and stochastic sampled) are discussed in the following subsections.

3.2.1. Greedy decisioning policy

The most intuitive and straightforward decisioning policy involves calculating the expected payoffs for all of the possible actions, and choosing the highest possible value - the so-called greedy policy [19]. However, such a method may easily lead the learning algorithm to ”get stuck” in a local optimal solution. An example is given in Figure 2, where a player fairly quickly learns which action is the unique best response and starts using it. However, once that its opponent catches up with this strategy and adapts, it will take significant time for the player’s expected payoff for the given action to drop below the other values, where in the meantime it will sustain significant payoff losses.

3.2.2. Stochastic sampled decisioning policy

A better approach may be obtained by using a stochastic sampled policy where, at each step, a randomly sampled action is taken with a probability \( p \). Sampling is performed by scaling the expected payoff value of each action to the minimum possible payoff for the game. For a minimum payoff \( P_{\text{MIN}} \) and
\( n \) choices with expected payoffs, \( P_1, \ldots, P_n \) the probability of choosing an action \( i \) is given as follows:

\[
P_i = \frac{P_i - P_{\text{MIN}}}{\sum_{k=1}^{n} P_k - P_{\text{MIN}}}
\]  

(7)

4. SIMULATION RESULTS

A game with the following parameters is observed: \( R_T = R_J = 10; \ H = 1; \ T_{\text{MAX}} = J_{\text{MAX}} = 2; \ n_f = 2 \). In the algorithm’s learning phase, it is enforced that every system state has been passed through exactly once. In this way, players’ learning vectors are initialized with the original state payoffs. Then, the game evolves on its own.

Fig. 3 shows comparison of the overall payoffs for three games: in all three, transmitter is playing the proposed Game Theory Optimal (GTO) strategy, whereas jammer is playing GTO in game 1, taking its decisions in every step randomly in game 2 (hence, regardless of the observations on the transmitter’s strategy), or always plays a fixed strategy \((C_i^f, f_s^f)= (2,2)\) in game 3 (i.e. it will always transmit at frequency 2 with power 2, again regardless of the observations). When both players are playing GTO, any deviation from the strategy would result in the decrease of the anticipated payoff for the deviating player.

Fig. 3: Comparison of overall payoffs for varying strategies of the jammer when transmitter plays GTO

Fig. 4 shows differences in the state transition probabilities for the transmitter as the game parameters change.

Fig. 4: Markov chain state transition probabilities for the transmitter as the game parameters change.

Future work will include reducing the action space of the MDP and - thus - the computational complexity of the algorithm by performing sampling in the learning process [20]. The game will also be extended to arbitrary number of jammers, including the case of cooperative jamming. Furthermore, performance of the adaptation of the proposed scheme will be tested using the real-life Software Defined Radio platforms [21].

Acknowledgements

This work was partially developed within the nSHIELD project (http://www.newshield.eu) co-funded by the ARTEMIS JOINT UNDERTAKING (Sub-programme SP6) focused on the research of SPD (Security, Privacy, Dependability) in the context of Embedded Systems.

This work was supported in part by the Erasmus Mundus joint Doctorate in interactive and Cognitive Environments, which is funded by the EACE Agency of the European Commission under EMJD ICE.
6. REFERENCES


