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 prospectives 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

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

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 spectrum 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, allowing it to create interference only at a single channel at a time. Furthermore, it is equipped with spectrum sensing capabilities [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 (channel hopping) [18], or transmitting at a higher power (power alteration).

3. GAME FORMULATION

Analyzing RF jamming and anti-jamming strategies is a complex 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 [1] iv) all the players in the game maintain their relative positions 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 every 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 successful transmission or jamming, as well as the costs of frequency hopping and the transmission cost. Payoff at the end of the step s for transmitter T is given as:

$$P_s^T(C_s^T, f_s^T, C_s^J, f_s^J) = R^T \cdot \alpha - H \cdot \beta - C_s^T, \quad (1)$$

[1] In real-life communication systems, identifying whether the signal is successfully jammed is more complex, and typically involves stochastic processing. For example, in analog voice communication systems, signal may be considered jammed if 30% or more of the transmitted voice messages are incomprehensible at the receiver side. Digital communication systems, on the other hand, exhibit a threshold effect, where there is a certain SINR below which BER greatly raises and the system performs poorly. Probability that the achieved SINR at the receiver is below this threshold may then be considered as measure of the jamming effectiveness.

Here, R^T is the reward for successful transmission, H is the fixed cost of hopping, C^T is the transmitter's current cost of transmission, f^T is the frequency that transmitter is currently using, α denotes event of successful transmission (2), and $\beta = 1$ if the transmitter decides to hop and $\beta = 0$ otherwise.

$$\alpha = \begin{cases} 1 & \text{if } C_s^T > C_s^J \text{ or } f_s^T \neq f_s^J \\ 0 & \text{if } C_s^T \leq C_s^J \text{ and } f_s^T = f_s^J \end{cases} \quad (2)$$

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

$$P_s^J(C_s^T, f_s^T, C_s^J, f_s^J) = R^J \cdot (1-\alpha) - H \cdot \gamma - C_s^J \quad (3)$$

R^J is jammer's reward for successful jamming, C_s^J is 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 $1 \leq C_s^T \leq T_{MAX}$; $1 \leq C_s^J \leq J_{MAX}$ with $T_{MAX} \leq J_{MAX}$.

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, transmission power and transmission frequency. In case that the players 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 mechanism. 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): keep and stay (KS), restart and change (RC), increase and stay (IS). The magnitude of the power increase ΔC^T is the minimum increase that the transmitter needs to invest in order to get the SINR at the receiver side over the threshold that guarantees a successful transmission. Correspondingly, increase of the power for the jammer relates to the minimum level of additional invested power required to ensure successful jamming on a given channel.

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

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

However, it is legitimate to expect that the learning mechanisms 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.

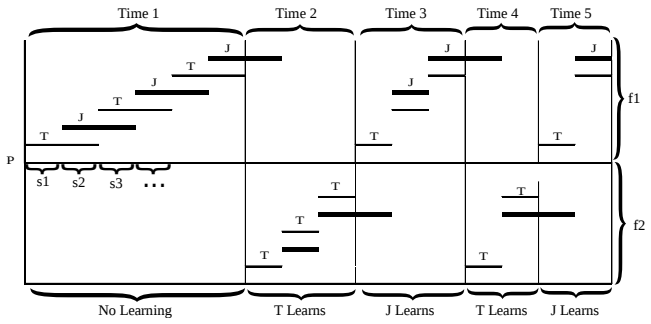


Fig. 1: Illustration of the arms race of the players' learning mechanisms.

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 ΔC^T (transmitter) or ΔC^J (jammer). This is repeated for as long as $C^T_{s+1} < H$ and $C^T_{s+1} \leq T_{MAX}$. 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 $\overline{P^T}$. For n_T possible discrete values of transmission powers and n_f channels in the system, cardinality of the vector is $|\overline{P^T}| = (n_T \cdot n_f)^2$. 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 $\overline{P^T}$ according to (6).

$$\overline{P^T_{s+1}(C^T_s, f^T_s, C^T_{s+1}, f^T_{s+1})} = \frac{\sum_{i=1}^{i=n_T} \sum_{j=1}^{j=n_f} N(C^T_{s+1}, f^T_{s+1}, i, j) P(C^T_{s+1}, f^T_{s+1}, i, j)}{\sum_{i=1}^{i=n_T} \sum_{j=1}^{j=n_f} N(C^T_{s+1}, f^T_{s+1}, i, j)}, \quad (6)$$

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 $\overline{P^J}$.

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 straight-forward 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.

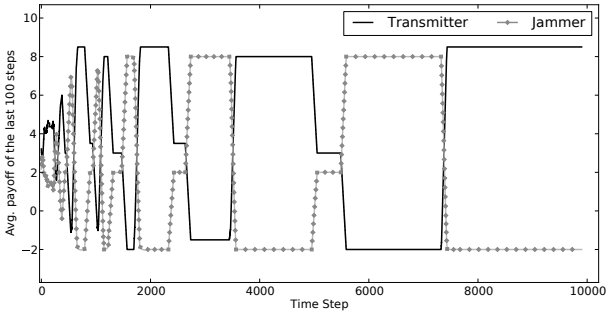


Fig. 2: Expected payoff over time for the greedy decisioning policy.

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_{MIN} and

n choices with expected payoffs, $P_1, \ldots, P_n$ the probability of choosing an action i is given as follows:

$$p_i = \frac{P_i - P_{MIN}}{\sum_{k=1}^{n} P_k - P_{MIN}} \tag{7}$$

4. SIMULATION RESULTS

A game with the following parameters is observed: $R_T = R_J = 10$; $H = 1$; $T_{MAX} = J_{MAX} = 2$; $n_f = 2$. 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_s^J, f_s^J)=(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.

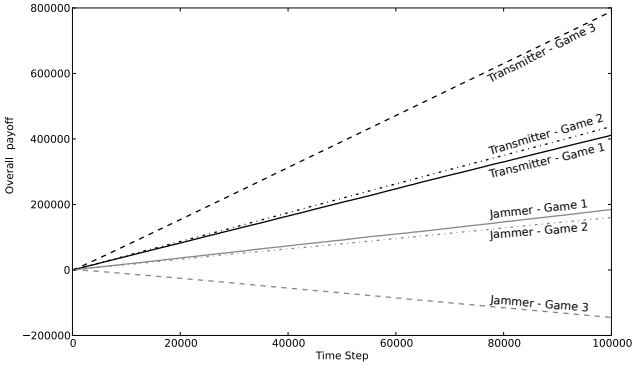


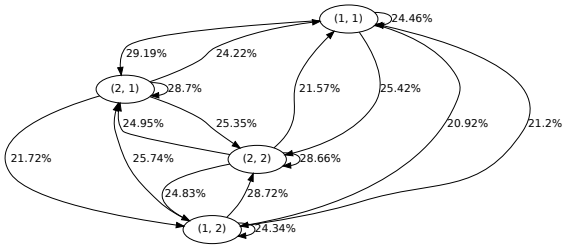
Fig. 3: Comparison of overall payoffs for varying strategies of the jammer when transmitter plays GTO

Fig. 4 shows differences in the state transition probabilities for the transmitter for differing game parameters. It can be seen how an increase in the hopping cost will directly influence transmitter's tendency towards hopping - for higher hopping cost (b), transmitter will be placing more value towards altering its transmission power or staying in the same state.

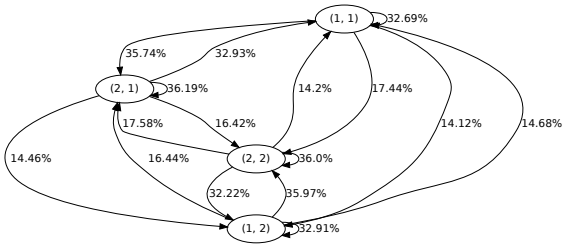
5. CONCLUSIONS AND FUTURE WORK

In the paper, we have modelled a Cognitive Radio jamming game between two players as a MDP. Increased action space included having both channel hopping and power alteration as anti-jamming strategies. Proposed learning algorithm based on fictitious play and stochastic sampled decisioning policy allowed for finding the game theory optimal solutions for both the jammer and the transmitter. Performance of the algorithm was supported by the simulation results.

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].



(a) $R_T = R_J = 10$; $H = 1$; $T_{MAX} = J_{MAX} = 2$; $n_f = 2$



(b) $R_T = R_J = 10$; $H = 8$; $T_{MAX} = J_{MAX} = 2$; $n_f = 2$

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

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

[1] J. Mitola and Jr. Maguire, G.Q., "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13–18, Aug 1999.

[2] K. Dabcevic, L. Marcenaro, and C. S. Regazzoni, "Security in cognitive radio networks," in *Evolution of Cognitive Networks and Self-Adaptive Communication Systems*, M. Louta T. D. Lagkas, P. Sarigiannidis and P. Chatzimisios, Eds., pp. 301–333. IGI Global, 2013.

[3] N.T. Nguyen, R. Zheng, and Z. Han, "On identifying primary user emulation attacks in cognitive radio systems using nonparametric bayesian classification," *Signal Processing, IEEE Transactions on*, vol. 60, no. 3, pp. 1432–1445, 2012.

[4] K.M. Borle, B. Chen, and W. Du, "A physical layer authentication scheme for countering primary user emulation attack," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 2935–2939.

[5] M. Abdelhakim, L. Zhang, J. Ren, and T. Li, "Cooperative sensing in cognitive networks under malicious attack," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 3004–3007.

[6] A.S. Rawat, P. Anand, H. Chen, and P.K. Varshney, "Countering byzantine attacks in cognitive radio networks," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, 2010, pp. 3098–3101.

[7] P. Tague, "Improving anti-jamming capability and increasing jamming impact with mobility control," in *Mobile Adhoc and Sensor Systems (MASS), 2010 IEEE 7th International Conference on*, 2010, pp. 501–506.

[8] R. Di Pietro and G. Oligeri, "Jamming mitigation in cognitive radio networks," *Network, IEEE*, vol. 27, no. 3, pp. 10–15, 2013.

[9] P. Morerio, K. Dabcevic, L. Marcenaro, and C.S. Regazzoni, "Distributed cognitive radio architecture with automatic frequency switching," in *Complexity in Engineering (COMPENG), 2012*, june 2012, pp. 1–4.

[10] R. Poisel, *Introduction to Communication Electronic Warfare Systems*, Artech House, Inc., Norwood, MA, USA, 2 edition, 2008.

[11] R. Poisel, *Modern Communications Jamming: Principles and Techniques*, Artech House intelligence and information operations series. Artech House, 2011.

[12] B. Wang, Y. Wu, K.J.R. Liu, and T.C. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 4, pp. 877–889, 2011.

[13] C. Chen, M. Song, C. Xin, and J. Backens, "A game-theoretical anti-jamming scheme for cognitive radio networks," *Network, IEEE*, vol. 27, no. 3, pp. 22–27, 2013.

[14] N. Buchbinder, L. Lewin-Eytan, I. Menache, J. Naor, and A. Orda, "Dynamic power allocation under arbitrary varying channels: an online approach," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 477–487, Apr. 2012.

[15] A. Garnaev, Y. Hayel, and E. Altman, "A bayesian jamming game in an ofdm wireless network," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012 10th International Symposium on*, 2012, pp. 41–48.

[16] J. Robinson, "An iterative method of solving a game," *Annals of Mathematics*, vol. 54, no. 2, pp. pp. 296–301, 1951.

[17] M. O. Mughal, L. Marcenaro, and C. S. Regazzoni, "Energy detection in multihop cooperative diversity networks: An analytical study," *International Journal of Distributed Sensor Networks*, vol. 2014, 2014.

[18] W. Xu, "Channel surfing: Defending wireless sensor networks from interference," in *Information Processing in Sensor Networks, 2007. IPSN 2007. 6th International Symposium on*, April 2007, pp. 499–508.

[19] K. Wang, Q. Liu, and L. Chen, "Optimality of greedy policy for a class of standard reward function of restless multi-armed bandit problem," *Signal Processing, IET*, vol. 6, no. 6, pp. 584–593, 2012.

[20] E. Sisikoglu, M.A. Epelman, and R.L. Smith, "A sampled fictitious play based learning algorithm for infinite horizon markov decision processes," in *Simulation Conference (WSC), Proceedings of the 2011 Winter*, 2011, pp. 4086–4097.

[21] K. Dabcevic, L. Marcenaro, and C. S. Regazzoni, "Spd-driven smart transmission layer based on a software defined radio test bed architecture," in *Proceedings of the 4th International Conference on Pervasive and Embedded Computing and Communication Systems*, 2014, pp. 219–230.