Adversarial Jamming and Catching Games over AWGN Channels with Mobile Players

Giovanni Perin*, Alessandro Buratto*, Nicolò M. Anselmi*, Shruti Wagle†, and Leonardo Badia*
* Dept. of Information Engineering (DEI), University of Padova, Italy
† Dept. of Information and Communication Technology, Manipal Institute of Technology, India
email: giovanni.perin.2@phd.unipd.it, {alessandro.buratto.1, nicolomaria.anselmi}@studenti.unipd.it,
shruti.wagle@learner.manipal.edu, leonardo.badia@unipd.it

***Abstract*—We consider a wireless jamming game played by a receiver and a malicious interferer that wants to disrupt communication. The outcome of the game depends on the characteristics of the wireless medium with distance-dependent path loss. The players can leverage this by changing their physical location. We consider both a static scenario where a position is kept forever by the players, and a dynamic one where players can change it over subsequent steps. We also include an optional feature for the receiver to *catch* the jammer when it is too close, which nullifies its jamming. The conclusion is that, in a mobile scenario, losing players see an improvement of their payoffs (or better, they cut their losses). As a result, we characterize the *Price of Mobility*, i.e., the benefit obtained thanks to the ability of changing position.**

***Index Terms*—Wireless communications; Jamming; AWGN; Game theory; Adversarial games; Zero-sum games.**

I. Introduction

THE SCENARIO of adversarial jamming [1] is a typical context where game theory is applied to wireless network communications [2]. As a sample reification of this idea, one can consider a legitimate receiver, which in the following will be referred to as the *user* (U). Its purpose is to receive data from a nearby wireless access point (AP), which we will consider to be non-controllable and stationary. The agent U is contrasted by an adversary called the *jammer* (J), acting to disrupt the wireless communication by increasing the noise, thereby lowering U's perceived signal-to-noise ratio (SNR) and, in turn, the achieved channel capacity.

In the present paper, we apply this rationale to the case of a "dumb" jammer guided by the sole goal of disturbing the communication between U and the AP [3]. This is reflected in an overall adversarial structure that just includes two players and defines their payoffs to obtain a zero-sum game [4] since the contrasting objectives of U and J are to maximize or minimize the communication capacity, respectively.

It is worthwhile noting that more sophisticated setups exist in the literature, where J may also be assumed to be equipped with more advanced capabilities and/or follows some stronger malicious intent [5]. For example, J can be interested in eavesdropping on the communication produced by the AP and/or be able to spoof it, either in the data or in some control information such as the global positioning [6]. For the sake of simplicity, we do not consider these security issues and just assume that J is only capable of raising the noise floor of U's intended communication. At the same time, player U is also assumed not to exploit any adaptive cognitive system; we remark that this would not solve the problem entirely but rather it would open the door to a more complex analysis within the same context, where the adversary J is also able to exploit the vulnerability of reconfigurable cognitive radios [7]. Finally, a different but related scenario is that of *friendly* jamming, which operates under the premises that U is a malicious user that wants to engage in some illegitimate communications and J is instead a network agent that is authorized and supported by the network to disrupt U's activity [8], [9].

Since jamming problems are usually limited to two agents, they are appealing from a modeling standpoint, and the literature is rich in game-theoretic investigations on this matter. Beyond the aforementioned references, we can also cite [10], which uses Bayesian games to analyze a type-uncertain situation about the role of the players, and [11], where we used a similar incomplete-information setup but to investigate the uncertainties related to the positions of the transmitters. Notably, one main difference between this last paper and the present contribution is that we have positions as a choice made by mobile players and not a trait captured as a randomly-determined type. Another recent trend of research, exemplified by [12], [13], involves the combination of game-theoretic models with explorations of the equilibria based on machine learning, which relates to our contribution since we use a Q-learning methodology to solve the dynamic games [14], [15].

Despite all these pre-existing contributions, we found out that the issue of mobility is scarcely addressed in the literature related to game theory and wireless jamming. Yet, the mobility of the nodes constitutes a fundamental aspect of wireless communications. Inspired by this motivation, we discuss the relatively unexplored scenario of an adversarial jamming game where nodes are mobile, and their position is chosen from a set of discrete locations over a grid. The channel additive white Gaussian noise (AWGN) and the jammer simply raise the noise floor depending on its proximity to the receiver. We investigate this system under both a static and a dynamic scenario, to imply just a choice of location that is kept forever in the former case, and an iterative update of position due to a mobility pattern, intentionally chosen by the node according to their selfish objectives in the latter.

Moreover, we consider two variations of the game. The first is a “pure jamming” in which the closer the nodes, the better for the jammer. However, we also consider a case where U can spot the jammer if they choose the same position. This results in the receiver getting a high utility; being aware of J’s presence, U can be assumed to physically intervene against J, or simply avoid its interference. This gives an interesting twist on the problem since it may be convenient for U to try “catching” J instead of fleeing from it.

The static scenario is analyzed from the point of view of a Nash equilibrium (NE). The dynamic scenario is considered in both a complete information context, for which we provide the subgame-perfect equilibrium (SPE), and an incomplete information one, where we discuss Q-learning strategies [14] and evaluate their effectiveness in achieving the SPEs. We will argue how the “pure jamming” scenario results in J following U as close as possible, while in the “catching” game U will take advantage of its best position.

By comparing the results, we are also able to evaluate the impact of mobility, meant as the option for the players to change their location over time. Mobility is found out to benefit the losing player, which is because, in the pure jamming game, U can benefit from being sometimes as close as possible to the AP, even if it is closely followed by J; on the other hand, in the catching game, J can choose its best location based on U’s best one. We also provide a quantitative assessment of the “Price of Mobility,” i.e., the ratio of U’s payoffs between the dynamic and the static scenario, which can be seen as U’s gain when players are allowed to change their positions, i.e., how much U is willing to pay for mobility.

The rest of this paper is organized as follows. In Section II we present the model for node placement and communication used in all the games. The different formulations of the game played between players J and U are discussed in III, first with a static setup that is subsequently extended to a dynamic game. In Section IV we show the numerical results, highlighting the Price of Mobility. Section V concludes the paper.

II. Communication Model and Setup

A. Spatial setup

We use a general wireless communication setting, not focusing on any particular modulation or technology. This means that we consider a plain AWGN channel between the involved communication parties. The communication system consists of three nodes: the first one is an AP, which is stationary and not controlled by a rational agent, it just passively transmits to the intended receiver. Then, we consider a receiver U and a jammer J, whose objectives are contrasting. The former wants to maximize the communication capacity over the AWGN channel, while the latter just causes a noise increase in the effort of choking the channel; in other words, J wants to minimize what U wants to maximize. Both U and J can choose their physical position as the action in a game.

For the sake of analytical tractability, the nodes are placed over the 3×3 square grid of Fig. 1. Nodes placed in a given square cell are assumed to be located in the middle. The AP is positioned in cell 5 and both J and U are forbidden to enter this cell, while they are free to choose any other one in $\{1, \dots, 4, 6, \dots, 9\}$. Even though this small grid may appear limited, it makes sense to restrict the choice of available positions to the ones close to the AP, since both players would be most interested in them: U to get the best possible signal from the AP, and J since being close to U allows for causing stronger interference. While the size of the grid is of relative importance since the system can be scaled, in the numerical evaluations we consider each square cell to be 10 m.

1	2	3
4	5	6
7	8	9

Fig. 1. 3×3 grid used for the proposed games.

B. Communication model and payoffs

We relate the payoff of the players to Shannon’s channel capacity achieved on the wireless channel. This in turn depends on the positions of the players and the AP. Without any jammer, the channel capacity depends on the SNR between the AP and U, as per the well-known relationships of the capacity of an AWGN channel. Thus, we can compute the SNR Γ as

$$\Gamma = \frac{P_T}{a(d_T) N_0 B} \tag{1}$$

where P_T is the transmitted power, $N_0 B$ is a noise term and $a(d_T)$ is the attenuation factor with respect to the distance between transmitter and receiver. We assume that $a(d_T)$ follows a power law with positive exponent α. Typically α is chosen between 2 and 4, or even wider, with 2 representing the free-space path loss. Formally,

$$a_{d_T} = K_0 \left(\frac{d_T}{d_0} \right)^{\alpha} \tag{2}$$

where K_0 is the attenuation experienced at reference distance d_0. In the following, we set d_0 to 1 m to simplify the formula. When a jammer is present, the noise term is increased by a jamming term, and the SNR translates to a signal-to-noise-plus-jamming ratio (SJNR) that we still denote as Γ. This can be computed by adding a term $P_J / a(d_J)$ to $N_0 B$ in (1). Once again, there is a dependence on the distance, in this case, the one between the jammer and the receiver, that we denote as d_J, thereby obtaining the following equation:

$$\Gamma = \frac{P_T}{a(d_T)[N_0 B + P_J / a(d_J)]} \tag{3}$$

The attenuation factor $a(d_J)$ follows the same relation described in (2) but with a different value of d_J. As a result, the SNR in the presence of a jammer can be written as

$$\Gamma(d_T, d_J) = \left[d_T^{\alpha} (K_1 + K_2 d_J^{-\alpha}) \right]^{-1} \tag{4}$$

expressing the dependence on both distances d_T and d_J. The terms K_1 and K_2 are suitable constants as found in [11].

Thus, we can compute the channel capacity C using the Shannon capacity formula for an AWGN channel

$$C = B \log_2 (1 + \Gamma) \tag{5}$$

where B is the available bandwidth and Γ is the previously computed SNR. If we take the spectral efficiency C/B of the channel in bit/s/Hz, we can cancel out the bandwidth and just consider a logarithmic dependence on Γ.

In the numerical evaluations, we consider $P_T = P_J$ and define K_1 and K_2 so that a user located at 10 m from the AP experiences an SNR of 40 dB and when the user stays in the previously described position and a jammer is present at 1 m distance from the user, the SNR drops to 20 dB.

III. Adversarial Jamming Games

To frame the interaction between U and J as a game, we make use of the following assumptions. U and J are players choosing a position in the grid, therefore their set of available actions is $\mathcal{A}_U = \mathcal{A}_J = \{1, \ldots, 4, 6, \ldots, 9\}$. Each interaction (i.e., a combined choice of positions for both U and J) results in a *utility* of player U, denoted as u_U as the spectral efficiency achieved, according to (5), which depends on the SJNR and therefore the combined action of both players. Given the adversarial nature of jamming in this context, we assume that J achieves utility u_J with the same absolute value of u_U but with the opposite sign; in other words, $u_J = -u_U$. This is a standard assumption in similar problems [4], [11] that allows representing the players as driven by contrasting objectives: U would like to maximize the overall achieved capacity, while J simply desires to minimize it, which would imply that maximum damage to the communication is caused.

Moreover, we consider two variations, as follows:

Pure jamming game: in this case, the rules apply as described above. It is worth noting that the best position for the jammer is always to choose the same chosen by U.

Catching game: in certain scenarios, it may not be sensible that U is oblivious to the presence of a jammer in the very same position. Possibly, U can physically see the jammer and at least avoid receiving (therefore saving energy) or even adopt some counteraction such as materially stop J from disrupting the communication. As a result, we consider also a scenario in which, whenever J chooses the same position as U, the jamming term is zero (as opposed to causing the maximum damage as per the pure jamming case).

A. Static scenario

We start from a static game of complete information $\mathcal{G} = \{\mathcal{A}_U, \mathcal{A}_J; u_U, u_J\}$. This is a game where the aforementioned interaction between U and J happens only once. Both players decide their position simultaneously and unbeknownst to each other, and their payoffs are evaluated. Formally speaking, players U and J have the same set of strategies coinciding with the choice of action. This game can be represented in normal form via an 8×8 bi-matrix.

Given the zero-sum trait of this game, we exploit standard game-theoretic results, such as the Nash theorem or the minimax theorem [2] to prove that the game must have a NE.

Remark 1: The game is a discoordination game [11], meaning that, beyond being zero-sum, it does not have a pure strategy that is strictly dominant for any player. Thus, its NEs must be in mixed strategies.

We also observe that this remark holds for both versions of a pure jamming and a catching game. Some considerations can also be derived from the symmetry of the payoff matrix. Intuitively, corner positions in the grid (i.e., odd-numbered cells such as 1, 3, 7, 9) have the same role for U, and so do side cells (i.e., even-numbered ones). Since even-numbered positions are closer to the AP, U may prefer them. In a pure jamming game, J also follows U. On the other hand, we expect that the NE of the catching game is less intuitive, since J would prefer to stay afar and there is an incentive for U to play the position where J is suspected to be.

B. Dynamic scenario

The dynamic scenario corresponds to the evolution of the static game described previously over time. Loosely speaking, we can consider it as a multi-stage repetition of the previously defined static interaction. However, for the sake of analytical tractability, we applied the following modifications.

First of all, players start the game in a given position, which is randomly determined with uniform probability among the eight possibilities. Then, players compute their stage payoffs and have the option to update their position with three available moves: (i) to move clock-wise; (ii) to move counterclockwise; (iii) to remain in the same location.

If we consider a complete-information dynamic game with infinite horizon, we can discuss more elaborate *strategies* of the players that characterize the gameplay by defining the action to play over multiple stages [2]. We can also exploit the concept of SPE, which expands over the idea of NE as a joint strategy choice of the players that result in a NE for all possible subgames.

For both pure jamming and catching games, it is possible to identify some effective strategies that lead to an SPE, as discussed in the following. Generally speaking, the trend is that one player is supposed to chase the other, depending on who is the winner if the two players meet in the same cell. As a result, J goes after U in the pure jamming case, while the opposite may happen in the catching game, depending on the value of the path loss exponent α. We can also formalize the following remarks

Remark 2: For the *pure jamming* game, the only SPE is found as U always moving away from the jammer, which in turn gives chase. Depending on the possible combinations of mutual positions, this ends up in both players moving around the AP clock-wise or counterclock-wise. Also, the two players can be in adjacent cells or with one empty cell in the middle.

Remark 3: For the *catching* game, a possible SPE implies that U chases J (which justifies the name) if J is close. Depending on the value of α, three possible SPE can be observed: i) the players go around the AP by keeping a distance of two cells; ii) U stops at one of the even cells, with J leaving one cell between U and itself, so that it is not caught; iii) U stops at an even cell, and J can choose one of the closest corners. The three cases happen for increasing α; specifically, the latter case happens when it is disadvantageous for U to transmit from an odd cell (i.e., a corner) even in absence of the jammer, because of the high impact of the distance from the AP.

A detailed analysis of these games, also proving the overall criteria for subgame-perfectness of these equilibria is out of the scope for the present paper and left as a further contribution. However, we also employed reinforcement learning (RL) to assess the performance of the described games in the case of *incomplete* information.

Indeed, the previously described dynamic game with complete information seems a bit unrealistic in that the players are aware at any time of their mutual positions. We may wonder what happens if the players just know of each other existence and also are informed about the propagation scenario and its parameters, but are ultimately in the dark about each other's position. At any rate, they can try to infer the position by the partial results of dynamic interaction, i.e., trying to estimate whether the opponent is located.

To this end, we consider a dynamic game with *incomplete* information, where players are aware of their position only (but not of the opponent's) and we use a RL framework is used to evaluate the average payoff obtained in such a context. Specifically, both players are aware of the AP position and the movement criteria described above. Also, for tractability reasons we consider a *finite* horizon, i.e., the game ends when the players find themselves in the same grid box. Both the *pure jamming* and the *catching* games are considered.

The used RL framework is based on a tabular version of the Q-learning algorithm (see, e.g., [14]), which aims at learning the value of taking action $a_t \in \mathcal{A}$ while in state $s_t \in \mathcal{S}$, observing the environment's responses. This is done via the recursive Bellman update

$$Q(a_t, s_t) = (1-\rho)\; Q\left(a_t, s_t\right) + \rho\left(r + \gamma \max_a Q\left(a, s\right)\right), \quad (6)$$

where $\rho \in [0, 1]$ weighs the importance of the most recent environment observations, $\gamma \in [0, 1]$ is the discount factor accounting for the relevance of the future, and r is the game's payoff in the current stage (also called *reward*). During execution, each players' (*agent* in the RL jargon) best policy is to decide for

$$a_{t+1} = \operatorname*{argmax}_a \; Q(a, s), \quad (7)$$

i.e., taking the action which maximizes the future discounted expected payoff. However, at the beginning the players must explore the environment's responses to their action to learn which is the actual optimal one in any given context. Therefore, the *exploration-exploitation* ϵ-greedy policy is added, selecting a random action with probability ϵ and the optimal one given by (7) otherwise [14]. The probability ϵ is chosen to be exponentially decaying in time, so that the exploration behavior is mostly concentrated at the beginning.

Tabular Q-learning is provably convergent to the optimal policy, given that a sufficient amount of time is provided to the agents for learning, and that the game can be modeled as a stationary Markov decision process (MDP). Stationarity does not hold in the considered scenario, as two players are learning their respective policies contemporarily [16]. A possible solution is to modify rule iii) augmenting the state dimension with some opponent's information, as done, for instance, in [15]. The addition of the opponent's position as a piece of state information is investigated in this work.

C. *The Price of Mobility*

In line with theoretical contributions in the field of game theory [5], the proposed setup and the comparison between the static and dynamic scenario allows us to define a metric that we call the *Price of Mobility*. It is defined as the ratio between the equilibrium payoff of player U in the dynamic and the static case. We remark that, since the game is zero-sum, such a payoff also represents the *value* of the game [2].

We expect that this Price of Mobility is a quantity greater than 1 in the pure jamming game, and lower than 1 in the catching one, since allowing the players for a position change is advantageous for the looser ones. Regarding the pure jamming game, although J can follow U closely, as discussed in Section III-B, U can benefit from being half of the time in the even cells, close to the AP, while never being prevented from transmitting, unlike what happens in the static scenario. On the other hand, in the catching game, U finds itself in a leading position, and J must be more conservative. With mobility, U is allowed to reach its favorite position, i.e., one of the even cells, but J can decide, even reactively, which is its best strategy knowing U's position.

IV. Numerical Results

We now present some numerical sample results. The NEs of the static game are computed through Gambit, a free software designed to numerically solve games. The evaluation of the SPEs in the dynamic game with complete information follows the same approach, adapted with a combinatorial evaluation of the different strategies as discussed previously.

For the incomplete information scenario, we developed a customized Q-learning framework in Python, with the following specifications. The updating coefficient $\rho = 0.9$ is set to a high value to favor the most recent observations since the environment is non-stationary. The discount factor is set to $\gamma = 0.99$, and 10 simulations are run and averaged to obtain the results. Each simulation lasts 5000 steps, and the last 1000 are considered for evaluation, with a residual exploration probability of $\epsilon = 0.01$. Out of the 10 simulations, the ones with the highest and lowest user's average payoff are discarded as outliers.

From these evaluations, we collect the value of the game (i.e., the utility of player U, since J's utility is just the opposite) for the three scenarios of static, dynamic with complete information, dynamic with incomplete information, and we plot it for the two scenarios of *pure jamming* and *catching* games, in Figs. 2 and 3, respectively. Both figures consider the path loss exponent α as the independent value. As a general result, all the curves exhibit a decreasing trend of U's payoff versus α, which is somehow to be expected given that the propagation conditions become less favorable and therefore the channel capacity is lower anyways. Also, the payoff in the catching game is generally higher as the jammer is prevented from moving too close to U.

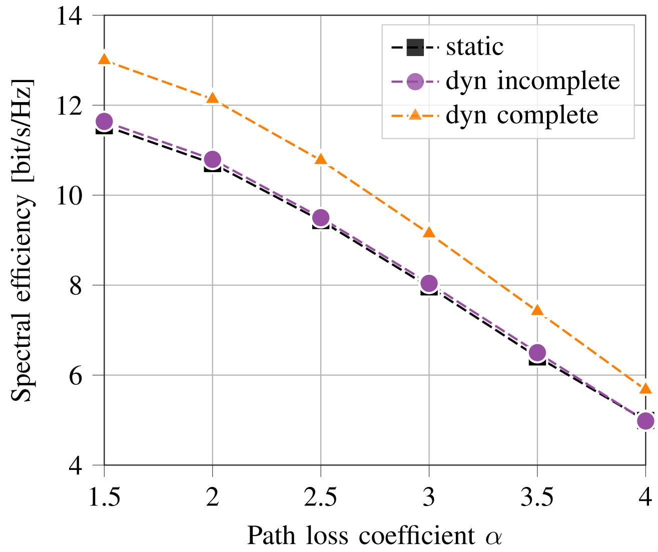

Fig. 2. User's payoff for the *pure jamming* game.

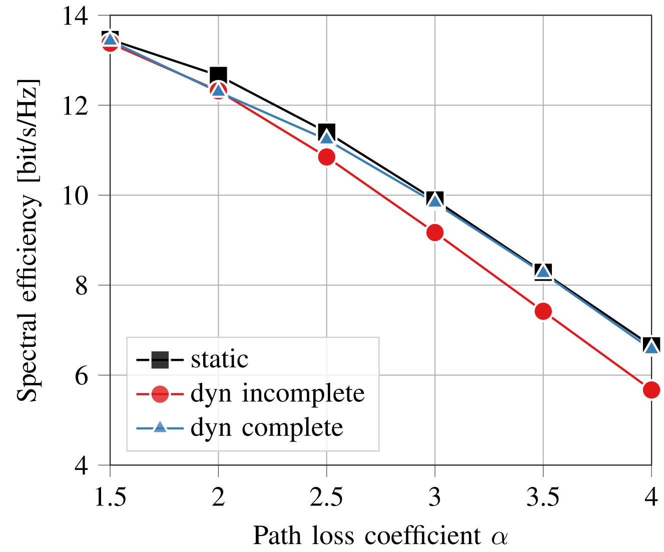

Fig. 3. User's payoff for the *catching* game.

A. Static game

In a static game, the results are consistent regardless of the value of α. In particular, the NE mixed strategy combines all available actions by the players, at least as long as $\alpha < 4$. When α reaches value 4 (or higher), player U starts playing only the even-numbered cells. This is because the path loss becomes so strong that it is not convenient for U to choose the corner cells that are further from the AP. Consequently, in the *pure jamming* game, the jammer tries to follow U and chooses the same cells, which results in only one NE where J just copies U's preferences. The case of the *catching* game is slightly more involuted, since J randomizes over all available cells even in the case of $\alpha = 4$ (since its goal is to be close to U but not in the same cell). From the theoretical point of view, we find four different NEs in this case, but this is just due to geometrical symmetry, as all these equilibria achieve the same payoff.

B. Dynamic game

The dynamic game is considered under both cases of complete and incomplete information.

The dynamic game with incomplete information where players do not know the opponent's position shows that the adopted *Q-learning* approach can recover a fair estimate of this information over repetitions of the gameplay. Regarding the *pure jamming* game in Fig. 2, it can be seen that there is a clear advantage for the user in moving with the dynamic infinite horizon game with complete information. This is because, this time, the user can benefit from being closer to the AP as an effect of mobility concerning the static solution. A slight improvement holds also for the Q-learning solution, where the lack of information and the problematic convergence of RL frameworks in adversarial contexts, prevent U from taking full advantage of mobility. There are residual occasions where the RL policy of player J can find U, stopping its transmissions. As mentioned before, the average spectral efficiency experienced by the user is generally higher in the *catching* game (Fig. 3). However, in this game, mobility produces a decrease in U's spectral efficiency, and the static solution represents in this case an upper bound. This produces an advantage for J, which is able, with its mobility, to react in the best way to U's best position choice. Specifically, in the case of complete information and with the chosen parameters, U always finds its best strategy in staying still in one of the even cells after reaching it. Concerning J, instead, it can place itself next to U as long as $\alpha \geq 2$, while it must leave an empty cell if $\alpha = 1.5$, for the reasons discussed in Section III-B. The *Q-learning* solution is again slightly suboptimal: instead of staying in its best position, U tries to catch the approaching J reactively, losing spectral efficiency while departing from even cells. Actually, the gap with the static solution increases with increasing α, since the impact of the distance is higher.

The addition of the opponent's position in the state information for the Q-learning algorithm is investigated in this section. A statistically significant difference in terms of the average payoff is not observed. Despite this, the player's behavior is different, reflecting a higher degree of consciousness of the process. Regarding the average episode length, i.e., the time between the beginning and the event corresponding to one of the players being caught by the other, it is expected the *catching* game to last more, as J's strategy is more conservative. This is actually observed in both information cases. However, when the opponent's position is unknown, both games last ~ 100 epochs, with an increase of 8% for the *catching* game. The magnitude order does not change for this game if the players know each other's position, but, with this knowledge, the *pure jamming* game lasts ~ 10 epochs only: J, knowing U's position, can find it rapidly, preventing it from transmitting. Therefore, a $10\times$ gain is observed regarding the episode length.

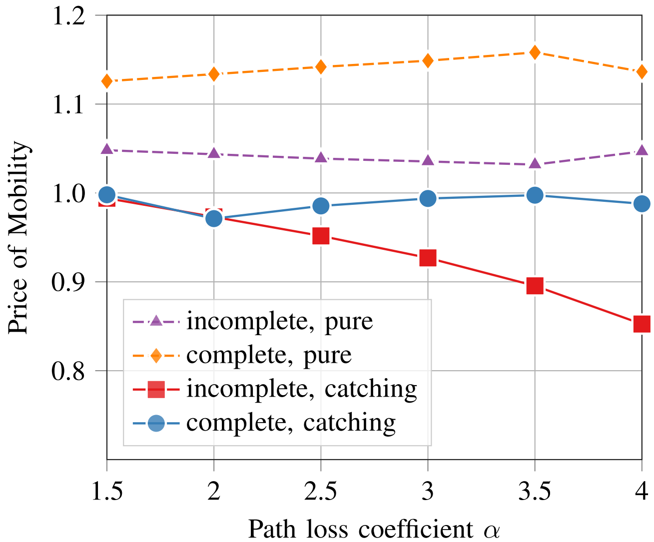

Fig. 4. The Price of Mobility, defined as the ratio between the static and the dynamic payoffs.

C. *The Price of Mobility*

We also report the Price of Mobility as defined previously. This is depicted in Fig. 4. As argued previously, the option to change position is generally found to be advantageous for the losing players in the respective games. Accordingly, the Price of Mobility is greater than 1 regarding the pure jamming game, and smaller than 1 otherwise.

Especially, the pure jamming game shows an almost constant Price of Mobility in α in the case of RL, with a gain of approximately 5%. Considering complete information, instead, the trend is increasing in α, as being close to the AP half of the time gets more advantageous as the impact of the distance increases. This is true except for $\alpha = 4$, where a change of static NE is observed. In the complete information case, the gain for U can be up to 15%.

The catching game, instead, exhibits a decreasing Price of Mobility in α for the incomplete information dynamic game, since, as discussed above, U keeps trying to chase J even when it is disadvantageous because of the increased distance from the AP. When, instead, complete information is given to both players, a minimum is observed for $\alpha = 2$, which is the point for which the SPE of the dynamic game changes. Beyond this point, J will be attached to U, while U knows that this is also its best possible condition. As expected, the Price of Mobility is strictly smaller than 1 (and very close to the unit for $\alpha = 3.5$), producing thus an advantage for J.

V. Conclusions and Future Work

We analyzed an adversarial game of receiver-versus-jammer, allowing the players to be mobile and choose their position according to different rules. We considered both a static and a dynamic version of the game, and we also included two different variants of "pure jamming" or "catching" games. We evaluate the Price of Mobility, defined as the gain encountered by the losing side when players are allowed to change their position over time.

An immediate future extension is the theoretical investigation of the structure of the SPEs in a dynamic context, also considering different grids and movement options. The study of possible mobility patterns defined a priori and their comparison in terms of overall efficiency appears intriguing.

Other future extensions of the present analysis involve a Bayesian investigation if the position of the nodes is uncertain. In our opinion, it would be particularly interesting to discuss how this can be ascertained, e.g., based on channel sensing paradigms or simply by updating the Bayesian prior after each iteration of the game. Another possible expansion of the present analysis would be to include incomplete knowledge about the actual presence of the jammer, or the extent of its adversarial role - for example, a detected node may either be a willing jammer or just an oblivious transmitter that is unaware of causing interference.

References

[1] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attacks and network defense policies in wireless sensor networks," in *Proc. IEEE INFOCOM*, 2007, pp. 1307–1315.

[2] L. A. DaSilva, H. Bogucka, and A. B. MacKenzie, "Game theory in wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 110–111, 2011.

[3] C. W. Commander, P. M. Pardalos, V. Ryabchenko, S. Uryasev, and G. Zrazhevsky, "The wireless network jamming problem," *J. Comb. Optim.*, vol. 2007, no. 14, pp. 481–498, 2007.

[4] X. Xu, K. Gao, X. Zheng, and I. Zhao, "A zero-sum game theoretic framework for jamming detection and avoidance in wireless sensor networks," in *Proc. IEEE CSIP*, 2012.

[5] M. Manshaei, Q. Zhu, T. Alpcan, T. Başar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Surveys*, vol. 45, no. 3, 2013, article no. 25.

[6] Y. Wang, J. M. Hao, W. P. Liu, , and X. Wang, "Dynamic evaluation of gnss spoofing and jamming efficacy based on game theory," *IEEE Access*, vol. 8, pp. 13 845–13 857, 2020.

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

[8] L. Badia and F. Gringoli, "A game of one/two strategic friendly jammers versus a malicious strategic node," *IEEE Netw. Lett.*, vol. 1, no. 1, pp. 6–9, 2019.

[9] J. Kim, P. K. Biswas, S. Bohacek, S. J. Mackey, S. Samoohi, and M. P. Patel, "Advanced protocols for the mitigation of friendly jamming in mobile ad-hoc networks," *J. Netw. Comp. Appl.*, vol. 1, no. 103037, 2021.

[10] A. Garnaev, A. P. Petropulu, W. Trappe, and H. V. Poor, "A jamming game with rival-type uncertainty," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5359–5372, 2020.

[11] M. Scalabrin, V. Vadori, A. V. Guglielmi, and L. Badia, "A zero-sum jamming game with incomplete position information in wireless scenarios," in *Proceedings of European Wireless 2015; 21th European Wireless Conference*. VDE, 2015, pp. 1–6.

[12] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab, and M. Guizani, "Iot anti-jamming strategy using game theory and neural network," in *Proc. IWCMC*. IEEE, 2020, pp. 770–776.

[13] Y. Arjoune, F. Salahdine, M. S. Islam, E. Ghribi, and N. Kaabouch, "A novel jamming attacks detection approach based on machine learning for wireless communication," in *Proc. ICOIN*. IEEE, 2020, pp. 459–464.

[14] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[15] A. Tampuu, T. Matiisen, D. Kodelja, I. Kuzovkin, K. Korjus, J. Aru, J. Aru, and R. Vicente, "Multiagent cooperation and competition with deep reinforcement learning," *PloS one*, vol. 12, no. 4, p. e0172395, 2017.

[16] L. Matignon, G. J. Laurent, and N. Le Fort-Piat, "Independent reinforcement learners in cooperative Markov games: a survey regarding coordination problems," *The Knowledge Engineering Review*, vol. 27, no. 1, p. 1–31, 2012.