Learning-Based Secure Spectrum Sharing for Intelligent IoT Networks

Amir Alipour-Fanid
Department of Computer Science and Information Technology
University of the District of Columbia
Washington, DC, USA
amir.alipourfanid@udc.edu

Monireh Dabaghchian
Department of Computer Science
Morgan State University
Baltimore, MD, USA
monireh.dabaghchian@morgan.edu

Long Jiao
Department of Electrical and Computer Engineering
University of Massachusetts Dartmouth
North Dartmouth, MA, USA
ljiao@umassd.edu

Kai Zeng
Department of Electrical and Computer Engineering
George Mason University
Fairfax, VA, USA
kzeng2@gmu.edu

2024 25th International Symposium on Quality Electronic Design (ISQED) | 979-8-3503-0927-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISQED60706.2024.10528684

***Abstract*—In intelligent IoT networks, an IoT user is capable of sensing the spectrum and learning from its observation to dynamically access the wireless channels without interfering with the primary user's signal. The network, however, is potentially subject to primary user emulation and jamming attacks. In the existing works, various attacks and defense mechanisms for spectrum sharing in IoT networks have been proposed. This paper systematically conducts a targeted survey of these efforts and proposes new approaches for future studies to strengthen the communication of IoT users. Our proposed methods involve the development of intelligent IoT devices that go beyond existing solutions, enabling them not only to share the spectrum with licensed users but also to effectively thwart potential attackers. First, considering practical aspects of imperfect spectrum sensing and delay, we propose to utilize online machine learning-based approaches to design spectrum sharing attack policies. We also investigate the attacker's channel observation/sensing capabilities to design attack policies using time-varying feedback graph models. Second, taking into account the IoT devices' practical characteristics of channel switching delay, we propose online learning-based channel access policies for optimal defense by the IoT device to guarantee the maximum network capacity. We then highlight future research directions, focusing on the defense of IoT devices against adaptive attackers. Finally, aided by concepts from intelligence and statistical factor analysis tools, we provide a workflow which can be utilized for devices' intelligence factors impact analysis on the defense performance.**

I. Introduction

Internet-of-Things (IoT) technology as a promising paradigm is envisioned to shift the future wireless communications and provide ubiquitous connections in many application areas such as smart city, smart home, smart vehicles, smart grid, smart farming, healthcare systems, etc. [1]–[3]. This projection, indeed, has now become close to reality by the recent emergence of intelligent IoT devices as a new technological design development which adds new capabilities such as sensing, learning, and reasoning to the IoT devices [4], [5].

However, the flourishing increase in the number of intelligent IoT devices causes explosive growth of demands for wireless spectrum bandwidth, and consequently spectrum shortage problem. To resolve the imminent spectrum shortage problem, Federal Communications Commission (FCC) has authorized opening spectrum bands (e.g., 3550-3700 MHz and TV white space) owned by licensed primary users to unlicensed secondary users when the licensed users are inactive [6].With this authorization, in wireless IoT networks, intelligent IoT users as unlicensed users can form the opportunistic spectrum sharing system to dynamically search and identify the unused portions of licensed spectrum (aka, spectrum hole or white space) to fully utilize that spectrum without adverse interference with the licensed users [7], [8].This opportunistic spectrum sharing mechanism results in increasing spectral efficiency in IoT networks. To enhance spectrum sharing mechanisms significantly, a recent development is the introduction of the Incumbent Informing Capability (IIC). This innovative approach aims to collaboratively, securely, and dynamically enhance opportunistic spectrum access within allocations primarily designated for federal government use [9]. Research in spectrum sharing for 5G IoT networks, focusing on 5G New Radio (NR) for IoT devices [10], and game theory-based spectrum sharing in industrial IoT networking [11], addresses distinct aspects of the IoT spectrum sharing problem.

Despite the significant advantages brought by opportunistic spectrum access mechanisms, IoT networks are vulnerable to threats posed by adversaries which aim to disrupt the communication between IoT users (see Fig. 1), and degrade the network capacity [12]–[14]. An adversary known as a primary user emulation attacker may attempt to occupy the spectrum hole [15]–[17], or as a jamming attacker, it may emit high power interference signal to interrupt the IoT users' communication [18]–[20].These threats justify the need for effective countermeasure mechanisms. The overall goal of this paper is to introduce a new class of secure, resilient, and efficient spectrum sharing frameworks for intelligent IoT devices with spectrum learning capabilities to circumvent the

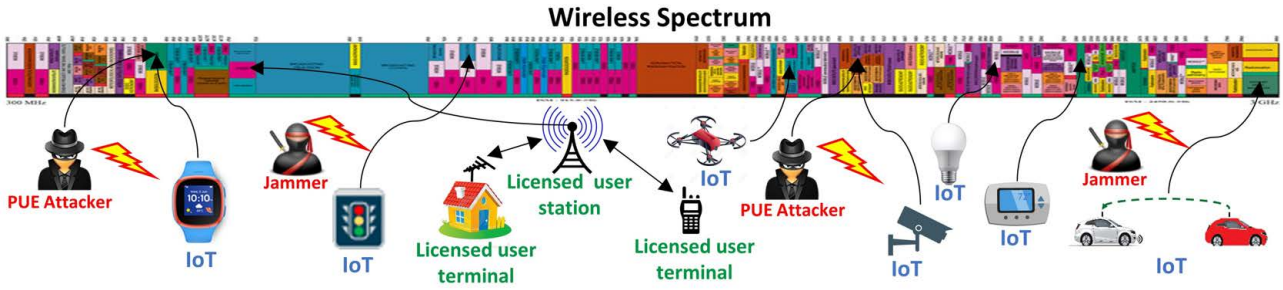

Fig. 1: Spectrum sharing of IoT networks in adversarial environments.

posed threats to maximize the network capacity.

Within the above goal, the first objective of this paper is to investigate the vulnerabilities of IoT wireless communication networks by proposing new online machine learning-based spectrum access attacking strategies. Better understanding of effective attacking strategies enables us to quantify and assess the severeness of attacks on the IoT networks which in turn sheds light on designing effective countermeasure schemes.

Spectrum sensing is a key enabling functionality for any online leaning-based attack policy in wireless spectrum sharing networks [21], [22]. Over time, by sensing a frequency channel, the attacker perceives whether there is an active IoT user communication on that channel. This information is utilized to update the attacker's belief, reasoning, and learning on the channel based on which it optimizes its attacking strategy. In practice, due to the attacker's inherent hardware imperfection and the computational complexities of the sensing algorithms, spectrum sensing introduces non-negligible delay and the sensing accuracy is imperfect which affects the attack performance [23], [24]. In our previous works [15], [25], along with the literature [26], a set of solutions for the optimal learning-based primary user emulation and jamming attacks have been proposed, respectively. However, these works have not considered the aforementioned practical aspects of spectrum sensing. In this paper, we aim to integrate the imperfect spectrum sensing and delay into the design paradigm of an effective attacking policy.

Attacker's *observation capability* is defined by the observation/sensing policy governing the attacks, as well as, the number of observations it makes. We study effective observation capabilities that lead to optimal attack mechanisms. Existing works consider deterministic policies with fixed number of observations [15], [25]–[28]. We study a random number channel observation policy using randomized time-varying feedback graphs. Following the proposed attacking policies in spectrum sharing IoT networks, we study a family of defense mechanisms for intelligent IoT devices to circumvent the attacker to achieve maximum network capacity. In our defense design paradigm, we take into account the inevitable limitations of IoT devices such as channel switching delay [29], [30], and the IoT devices' intelligence capabilities [4], [31], [32].

We consider an intelligent IoT device which aims to defend against an attacker with no prior knowledge about the attacker's attack policy nor of the channel state information such as noise and fading. The defense strategy employed by the IoT device will be governed by online machine learning approaches for effective sequential frequency channel selection and data transmission. In practice, the IoT device incurs a channel switching delay when switching from a certain frequency channel to a different channel [29], [30], [33]. The channel switching delay which leads to throughput loss is due to the elapsed time between detaching from the current operation frequency and resettling on another channel for data communication [34]. In the existing works, several defense strategies have been designed without consideration of switching delay [20], [26], [35]–[37]. We formulate an online learning-based defense mechanism for the intelligent IoT device to maximize network capacity while minimizing channel switching delay.

When both the attacker and IoT device apply learning-based frameworks to access the channels, the two agents become adaptive to each other's strategy [38]. For this setting, we study the asymptotic network capacity that can be achieved (degraded) by the IoT device (attacker). In the literature, the outcome of two adaptive opponents when both of the agents apply a learning-based strategy has been studied [26]. However, the authors assumed fixed and time-invariable channel statistics. We characterize the wireless channels' inherent conditions with a time-varying stochastic process and integrate its effects into the problem formulation. With this modeling, our setting forms an online repeated stochastic game between the IoT device and the attacker for which we illustrate an equilibrium under certain assumption.

Sophisticated intelligence or cognitive capabilities such as sensing, reasoning, and learning are essential for IoT devices to cope with the uncertainties of spectrum environment. Being able to quantitatively measure the intelligence capabilities of IoT devices enables us to design and deploy efficient and resilient IoT devices (see Fig. 2). The research on the intelligence measure of IoT devices is in its early stages. Several works aimed to study this problem [39]–[41]; however, cognitive factors extraction and their quantitative analysis are missing in these works. In our previous work, we proposed a data-driven methodology to quantitatively measure intelligence factors [31], [32] using statistical factor analysis [42], [43]. In this paper, we propose to characterize the IoT device's robustness as a function of its intelligence factors in adversarial wireless communication networks.

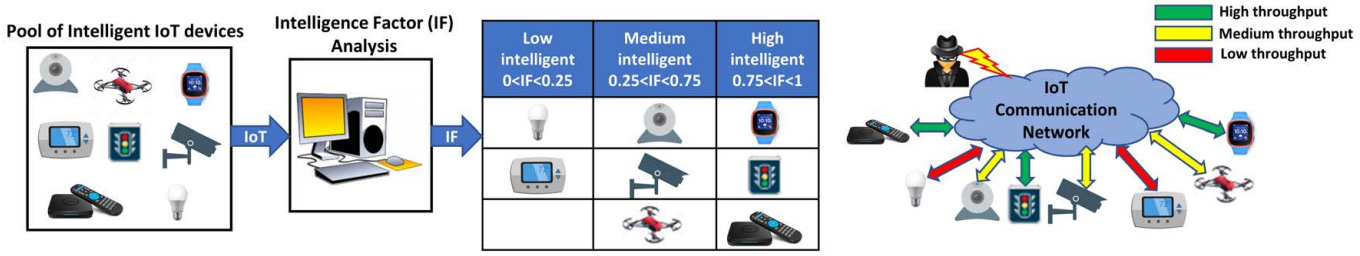

Fig. 2: Intelligence factor analysis of various IoT devices with different cognitive capabilities.

II. Background Information

Sharing spectrum with legacy systems has attracted intensive research during the past decade [44]–[46]. In order to resolve the imminent spectrum shortage problem, cognitive radios have been emerged as a key enabling technology for dynamic spectrum access to improve spectrum efficiency and guarantee the unharmful coexistence with the legacy systems [44], [45], [47], [48]. Although the cognitive radio concept was born with the core idea of realizing "cognition" [49], the research on measuring cognitive radios' cognitive capabilities or intelligence is largely open. There are various works on studying and evaluating cognitive radios' performance [39], [40], [50], [51]. In our previously work [31], motivated by the Cattell-Horn-Carroll (CHC) intelligence model [52], [53], we proposed a data-driven methodology to model and measure the cognitive capabilities of cognitive radios based on factor analysis [42], [43], [54]. The complex and uncertain spectrum environment makes IoT users' spectrum sharing extremely challenging. The uncertainty may come from the inherent nature of the IoT communication system such as fluctuations in wireless signal propagation and the legacy system activities, or it may be generated due to the presence of adversaries which aim to maliciously degrade the system performance.

Existing works study security of spectrum sharing networks by introducing various types of attacks including jamming and primary user emulation attacks [15], [16], [18], [19]. There are several works formulating jamming attacks and anti-jamming strategies as online learning problems [20], [26], [35], [36]. Optimal primary user emulation attack strategy has been also addressed in our previous work [15], [25] where the problem has been formulated as an online machine learning problem in the setting of adversarial Multi-Armed Bandits (MAB) [55].

Multi-armed bandits is one of the most fundamental online learning problems, wherein, at each round a player chooses an action out of K available actions and observes the reward associated with the chosen arm. The reward may either be stochastic or adversarial (aka, non-stochastic). Several real-world problems, especially those that involve sequential decision making, can be posed as multi-armed bandit problem. These include clinical trials, online advertisement, routing in communication networks, spectrum sharing in cognitive radio wireless communication networks, personalized matching, and many others. There are many works in the literature on MAB [56]–[60]. The popular EXP3 algorithm was proposed by [55], [57], and was inspired by prior work on weighted majority algorithm [61] and Hedge algorithm [62]. Online learning algorithms are efficient in time and space complexity where they best fit to address the problem of secure spectrum sharing in IoT networks. Some other multi-armed bandits work related to our paper are, MAB with feedback graph [60], restless MAB [63], distributed stochastic online learning [64], and online learning against an adaptive adversary [59].

III. Secure Spectrum Sharing for Intelligent IoT Networks

In this section, we thoroughly study the problem of secure spectrum sharing and reliable communication in the intelligent IoT networks.

A. Learning-based Attack Policy in IoT Networks

We investigate the IoT wireless communication networks security threats by designing and developing smart learning-based dynamic attacking strategies. In our attacking strategy design paradigm, we take into account the practical aspects of an attacker including imperfect sensing, sensing delay, and various observation capabilities.

1) Online Learning-based Attack Policy with Imperfect Sensing and Delay

For learning, the attacker senses the frequency channels and employs detection techniques such as matched filter, and energy detection to identify any IoT user communication on the channels [65]–[67]. However, imperfect sensing and sensing delay are two of the indispensable drawbacks of practical sensing systems which affect the attacker's learning quality, and efficiency. These shortcomings are due to the inherent limitations of device hardware, and computational complexities of sensing algorithms. In our previous works, we have considered prefect channel sensing for a learning-based attacker [15], [25]. However, in this paper, we approach the problem by integrating sensing delay and accuracy into the learning process of the attacker.

We model the attack policy with adversarial multi-armed bandits where the rewards (i.e., degraded throughput) are delayed and imperfect. Our problem best fits to this model as it does not require the attacker to have prior knowledge about the IoT users' activity on the channels. Based on our model, at each time, the attacker chooses one channel to launch an attack seeking to minimize the network capacity, and another channel to imperfectly sense. Through this imperfect observation, the

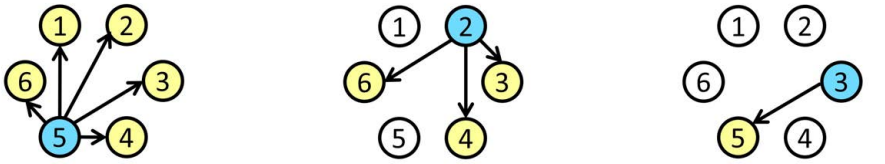

Fig. 3: Time-varying feedback graph, $K=6$ channels (blue:attacked, yellow:observed, white:not attacked and observed channels).

attacker either receives a delayed reward equivalent to the normalized degraded throughput or zero, depending on whether it identifies an IoT user on the channel or not. Within this policy, we model the sensing accuracy as a Bernoulli process with its parameter indicating the success rate of actual reward observation. Sensing delay is as well integrated in the model based on which the attacker will observe and update its policy with some delay. Note that existing frameworks of delayed reward in multi-armed bandits [68], [69] cannot be applied to solve our problem as they are designed for stochastic settings, whereas our problem setting is non-stochastic. This non-stochasticity is due to unknown IoT and licensed users activities, as well as, the attacker's attack pattern.

Regret is the metric defined for performance evaluation of online learning-based algorithms where it measures the difference between the accumulated reward achieved by the proposed algorithm and that of the optimal static policy in hindsight [55]. Obtaining a regret with a sublinear order in time T (i.e., $\lim_{T\to+\infty}\frac{R(T)}{T}=0$), indicates the attacker converges to attacking the channels with more active IoT users, consequently being effective to degrade the network capacity. The performance of the proposed attack policy can be formulated and measured by its regret as follows:

$$R(T)=\max_{j\subset[K]}\sum_{t=1}^{T}x_j(t)-\sum_{t=1}^{T}\left(1-p_{i_{t-\tau}}(t-\tau)\right)x_{i_{t-\tau}}(t-\tau), \tag{1}$$

where we divide the spectrum bandwidth into K frequency channels $[K]:=\{1,2,\ldots,K\}$. The sensing random delay is denoted by τ. $i_t\in[K]$ denotes the channel index chosen by the attacker at time t. The observed delayed reward is denoted by $(1-p_{i_{t-\tau}}(t))x_{i_{t-\tau}}(t-\tau)\in[0,1]$, with $i_{t-\tau}$, $x_{i_{t-\tau}}$, and $p_{i_{t-\tau}}(t-\tau)$ indicating the attacked channel, actual reward, and the probability of the actual reward observation, respectively. The normalized reward will characterize the degraded network capacity due to the attacker's imperfect sensing and delay. This analysis indicates that compared to the prefect situation, both imperfect sensing and delay inflate the regret.

2) Random Observation Policy with Time-varying Feedback Graphs for Learning-based Attack Policy

For effective learning, the attacker needs to make enough observation/sensing on all the channels. On the other hand, the attacker's learning rate does not necessarily improve by merely making more observations; rather it depends on the observation policy, as well. Modeling with feedback graphs, the attack performance may be optimal, sub-optimal, and even it may not learn at all, depending on the observation policy [60]. Unlike our previous work [15] which is based on a single sensing policy with a fixed number of observations, in this subsection, we investigate the random number of observations along with various observation policies governing the attack policy to design an effective and efficient attacker.

We propose an observation policy for the attacker based on time-varying feedback graphs. According to this policy, the attacker dynamically selects an observation policy, as well as a random number of m channels to observe, as shown in Fig. 3. In the beginning, since the attacker has no prior knowledge about the IoT users' activity, it leverages its maximum observation capability to observe as many channels as possible. However, over time, as the attacker learns the IoT users' activity, the attacker will observe fewer channels based on the policy such that the attacker will converges to observing a single channel uniformly at random. The proposed policy eliminates the redundant and unnecessary channel observations which is the case with fixed number of observations. Regarding observation policy, we propose to consider various types of policies such as uniformly at random, round-robin, bandit, loopless clique, etc., and utilize them as expert advise by assigning weights to each policy which will be updated at each time. The proposed policy provides a unified and efficient observation strategy for IoT spectrum sharing networks. This encompasses both the incorporation of random observations and the application of expert advice, resulting in an observation policy designed to optimize attack performance.

B. Learning-based Defense Mechanisms in IoT Networks

In this subsection, we study learning-based defense mechanisms by an IoT user considering its practical aspects including channel switching delay and the performance of an IoT device against an adaptive learning-based attacker. Furthermore, we investigate the intelligence capabilities of IoT devices and their impact on the resiliency of defense mechanisms.

1) Online Learning-based Defense Policy for IoT with Channel Switching Delay

A learning-based IoT device dynamically selects various channels to evade the attacker and transmit data on the higher throughput channels to maximize the network capacity [70], [71]. Due to hardware limitations and imperfections, switching from a certain frequency channel to another incurs an overhead in terms of delay while the radio takes time to actuate and settle [29]. This delay results in throughput loss and ultimately network capacity degradation which is non-negligible

in practice [30], [34]. Our goal is to design an effective learning-based defense policy for the IoT user that strikes an optimal balance between throughput maximization and channel switching delay minimization.

We model the problem as an adversarial multi-armed bandits with switching costs wherein not only it requires a careful trade-off between "exploitation" and "exploration" for effective defense, but also to account for channel switching delay. We adopt the method proposed in our previous work [8], [72], a randomized switching policy which follows a stochastic Bernoulli process. In this method, at each time, the IoT user will choose to stay on the same channel with probability $1-\delta(t)$ and it will switch with probability $\delta(t)$. The parameter of the Bernoulli distribution depends on the number of channels K, and should be decaying with time as $t^{-\alpha}$. The choice of α is crucial – a slow decaying $\delta(t)$ would allow frequent switching and help with exploration, at the expense of potentially not exploiting high throughput channels and incurring additional switching delays. On the other hand, a fast decaying $\delta(t)$ may hurt exploration and, therefore, overall throughput by resulting in a poor defense policy.

We define the IoT user's reward (normalized throughput $v \in [0,1]$) on channel i at time t as follows:

$$x_i(t) = \begin{cases} v, & \textit{if no attack on channel } i, \\ 0, & \textit{if channel } i \textit{ is under attack}. \end{cases} \tag{2}$$

The IoT user will incur a normalized throughput loss $c(t)$ for switching between channels over two consecutive times. We define the regret of the IoT user with channel switching delay after T rounds, as follows:

$$R(T) = \max_{j \subset [K]} \sum_{t=1}^{T} x_j(t) - \left(\sum_{t=1}^{T} x_{i_t}(t) - \sum_{t=1}^{T} c(t)\ \mathbb{1}_{\{i_t \neq i_{t-1}\}} \right), \tag{3}$$

where $i_t \in [K]$ indicates the channel index chosen for attack. We evaluated the performance of the proposed technique in our previous work [8], [72] by conducting theoretical analysis on its regret lower and upped bounds. The proposed algorithm to solve this problem is minimax optimal if the results match. We found $\alpha = 1/3$, such that δ_t proportional to $t^{-\alpha}$, will offer an optimal trade-off between exploration and exploitation yielding optimal defense policy [72], [73]. We proposed an online learning algorithms for IoT with and without channel switching costs, where their regret performances are proved sublinear order-optimal in time as $T^{2/3}$ and $T^{1/2}$, respectively, offering throughput-optimal for IoT spectrum sharing network. In addition, we provided numerical analysis, to validate the theoretical analysis. Our analysis can be adopted to measure the IoT device's throughput under various artificially injected attack signals over multi channels and compare the real-world performance results of defense with and without channel switching delays on the throughput loss. This provides an optimal spectrum access policy for the IoT users in a practical adversarial IoT network which maximizes the network capacity.

2) Stochastic Game-Theoretic Analysis of Learning-based IoT user and Attacker

Both IoT user and the attacker apply learning-based spectrum access policies which makes them adaptive to each other's strategy. We are interested in computing the asymptotic degraded throughput by the attacker to study any possible equilibrium between the IoT user and the attacker. Wang *et al.* [26] studied the outcome of two adaptive opponents and showed they achieve Nash equilibrium. This analysis has been done based on the assumption of ideal wireless communication channels with fixed time-invariant channel states; whereas, in practice the channel state information may follow any unknown stochastic process. The stochasticity of the channels introduces non-trivial new challenges in analyzing and deriving the equilibrium.

To approach this problem, we propose to utilize techniques from both stochastic optimization [74] and game theory [38], [75] to form an *online repeated two-player zero-sum stochastic game* between the IoT user and attacker. We define the game by an stochastic $K \times K$ payoff matrix denoted by $\mathbf{G(t)}$. The IoT user computes a mixed channel selection strategy according to probability vector of $\mathbf{p}$ to choose channel i_t for data transmission. The attacker, as well computes a probability vector of $\mathbf{z}$ as its mixed channel selection strategy to choose channel j_t to launch the attack signal. The IoT user then gains the quantity $\mathbf{G}_{i_t j_t}(t)$, while the attacker loses the same quantity. In this game, the IoT user aims to maximize its expected total gain $\overline{r}(\mathbf{p}, \mathbf{z}, t) = \mathbf{p}^T \mathbf{G(t)} \mathbf{z}$, while the attacker aims to minimize its expected total loss. Due to the stochastic quantities in the payoff matrix, the *game value* $V(t)$ will be an stochastic process expressed as

$$V(t) = \max_{\mathbf{p}} \min_{\mathbf{z}} \overline{r}(\mathbf{z}, \mathbf{p}, t) = \min_{\mathbf{p}} \max_{\mathbf{z}} \overline{r}(\mathbf{p}, \mathbf{z}, t). \tag{4}$$

To evaluate the IoT user and the attacker's performance, we propose to utilize stochastic simulation-based optimization methods. The method requires extensive simulations to compute the asymptotic achievable (degraded) network capacity by the IoT user (attacker) to find any possible equilibrium stationary strategy for various models of stochastic channels with an acceptable confidence interval. This study enables us to evaluate the asymptotic network capacity, and subsequently to obtain a deeper insight on the defense mechanism performance when the IoT user faces an adaptive opponent. In the proposed game model if the channel states are considered to be fixed $\mathbf{G(t)} = 1$, then the game will recover the results of the well-known online repeated zero-sum game where the empirical distribution of channel selection for the IoT user and attacker will converge to the uniform distribution over K channels (i.e., the Nash equilibrium stationary strategies $\mathbf{p} = \mathbf{z} = \frac{1}{K}$).

3) IoT Intelligence Capabilities in Defense Mechanisms

Intelligent IoT devices are equipped with various cognitive capabilities including channel access policy, number of sensors, sensing accuracy, memory, power resources, computational complexity, processing speed, etc., where each of them plays a key role in the IoT device's adaptability and resiliency against

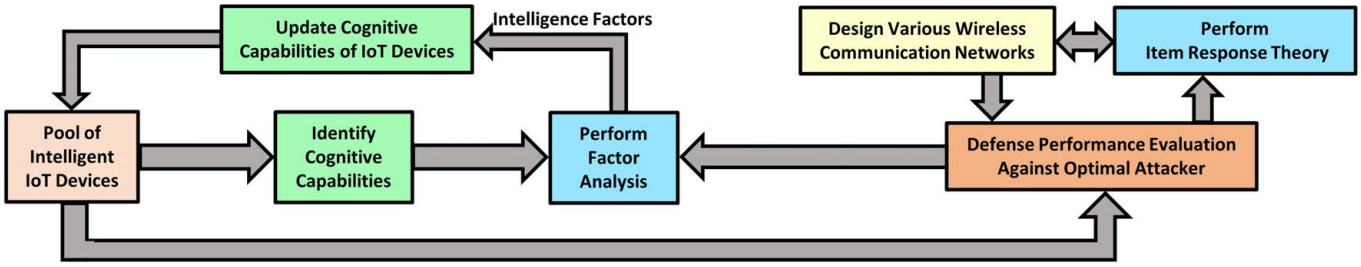

Fig. 4: Proposed workflow: Computing intelligence factors to guarantee the minimum required network capacity.

the attacker. We have previously investigated extraction of the intelligence factors of cognitive radios and the cognitive load associated with each intelligence factor [31], [32]. In this paper we provide a workflow which can be utilized to construct a relationship between the intelligence factors and the cognitive IoT device's resiliency in adversarial environment.

Fig. 4 demonstrates the workflow of the proposed method. According to this workflow, we propose to first identify the cognitive capabilities of a pool of intelligent IoT devices and derive their intelligence factors (latent factors) by applying statistical factor analysis method [31]. To derive these factors, we suggest to design various environments modeling different wireless communication networks and apply Item Response Theory [76] to quantify the level of hardness of each environment. Next, the performance of each IoT device against the optimal attacker is assessed by measuring key parameters such as throughput, delay, and interference level in the designated environments. Subsequently, based on the outcomes, the intelligence capabilities of the IoT devices are updated, and the same process is repeated until the minimum required network capacity is attained. Validation of this methodology is a part of our future work. We aim to generate a large pool of artificially intelligent IoT devices with various cognitive capabilities, operating within diverse wireless communication environments vulnerable to attacks. Then, extensive numerical analysis and simulations employing statistical data analysis techniques can be conducted [43], [54], [76]. The primary analysis will utilize SPSS software [54] for factor analysis, SAS software [77] for item response theory analysis, as well as AMPL and CVX open-source optimization toolboxes [78]–[80] to address optimization problems.

IV. Conclusion

We systematically surveyed spectrum sharing IoT networks and introduced an innovative set of strategies for future studies to improve the security of spectrum sharing within the IoT networks. By leveraging multidisciplinary approaches that integrate online machine learning, stochastic optimization, and game theory, these strategies offer robust communication among IoT users in wireless networks. We introduced two spectrum sharing attack policies, considering practical aspects of IoT device capabilities. Three defense mechanisms were explored, employing various methods such as utilizing online learning with channel switching costs, implementing a game-theoretic approach between IoT users and attackers, and establishing a framework for the intelligence capabilities of IoT devices to enhance defense mechanisms. As the demand for interconnected devices continues to grow, our proposed policies present a strategic and comprehensive solution to address the challenges associated with spectrum sharing. This paves the way for more resilient and secure IoT communication systems in the future.

V. Acknowledgment

This work was supported in part by the US National Science Foundation (NSF) Research Initiation Award under Grant No. 2100804, NSF Institute for Trustworthy AI in Law and Society (TRAILS) under Award No. 2229885, NSF Networking Technology and Systems (NeTS) Program with project No. 2131507, Microsoft Research Award, and Virginia Commonwealth Cyber Initiative (CCI), an investment in the advancement of cyber R&D, innovation, and workforce development.

References

[1] S. Balaji, K. Nathani, and R. Santhakumar, "IoT technology, applications and challenges: A contemporary survey," in *Wireless Personal Communications*, April 2019, p. 363–388.

[2] J. Ding, M. Nemati, C. Ranaweera, and J. Choi, "IoT connectivity technologies and applications: A survey," *IEEE Access*, vol. 8, pp. 67 646–67 673, 2020.

[3] Q. F. Hassan, *Internet of Things Applications for Agriculture*. IEEE Press Wiley, 2018, pp. 507–528.

[4] P. K. D. Pramanik, S. Pal, and P. Choudhury, *Beyond Automation: The Cognitive IoT. Artificial Intelligence Brings Sense to the Internet of Things*. Cham: Springer International Publishing, 2018, pp. 1–37.

[5] M. Kwon, J. Lee, and H. Park, "Intelligent IoT connectivity: Deep reinforcement learning approach," *IEEE Sensors Journal*, vol. 20, no. 5, pp. 2782–2791, 2020.

[6] Federal Communications Commission (FCC), "Report and order and second further notice of proposed rule making, 15-47 gn Docket no. 12-354," 2015.

[7] G. Caso, L. De Nardis, R. Thobaben, and M.-G. Di Benedetto, *Opportunistic Spectrum Sharing and White Space Access*. John Wiley & Sons, Ltd, 2015, ch. 7, pp. 143–165.

[8] A. Alipour-Fanid, M. Dabaghchian, R. Arora, and K. Zeng, "Multiuser scheduling in centralized cognitive radio networks: A multi-armed bandit approach," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 1074–1091, 2022.

[9] M. DiFrancisco, E. Drocella, P. Ransom, and C. Cooper, "Incumbent informing capability (IIC) for time-based spectrum sharing," *The 48th Research Conference on Communication, Information and Internet Policy*, 2020. [Online]. Available: http://dx.doi.org/10.2139/ssrn.3748483

[10] R. Saleem, W. Ni, and M. Ikram, "Reinforcement learning-based unlicensed spectrum sharing for iot devices of 5g new radio," in *2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, 2022, pp. 191–196.

[11] C. Wu, J. Sheng, Y. Wang, and B. Ai, "Game-theory-based spectrum sharing of industrial iot networking in high-speed railway heterogeneous communication system," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2023.

[12] F. Afghah, A. Shamsoshoara, L. L. Njilla, and C. A. Kamhoua, *Cooperative Spectrum Sharing and Trust Management in IoT Networks*. Wiley-IEEE Press, 2020, pp. 79–109.

[13] N. Wang, W. Li, A. Alipour-Fanid, M. Dabaghchian, and K. Zeng, "Compressed-sensing-based pilot contamination attack detection for NOMA-IoT communications," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7764–7772, 2020.

[14] N. Wang, W. Li, A. Alipour-Fanid, L. Jiao, M. Dabaghchian, and K. Zeng, "Pilot contamination attack detection for 5G mmwave grant-free IoT networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 658–670, 2021.

[15] M. Dabaghchian, A. Alipour-Fanid, K. Zeng, Q. Wang, and P. Auer, "Online learning with randomized feedback graphs for optimal PUE attacks in cognitive radio networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 5, pp. 2268–2281, 2018.

[16] I. Gupta and O. P. Sahu, "An overview of primary user emulation attack in cognitive radio networks," in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, 2018, pp. 27–31.

[17] H. Li and Z. Han, "Dogfight in spectrum: Combating primary user emulation attacks in cognitive radio systems, part I: Known channel statistics," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 11, pp. 3566–3577, November 2010.

[18] N. Namvar, W. Saad, N. Bahadori, and B. Kelley, "Jamming in the Internet of Things: A game-theoretic perspective," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.

[19] W. Wang, S. Bhattacharjee, M. Chatterjee, and K. Kwiat, "Collaborative jamming and collaborative defense in cognitive radio networks," *Pervasive and Mobile Computing*, vol. 9, no. 4, pp. 572 – 587, 2013.

[20] Q. Wang, K. Ren, and P. Ning, "Anti-jamming communication in cognitive radio networks with unknown channel statistics," in *19th IEEE International Conference on Network Protocols*, 2011, pp. 393–402.

[21] Y.-C. Liang, *Spectrum Sensing Theories and Methods*. Singapore: Springer Singapore, 2020, pp. 41–85.

[22] D. Lee and H. Wu, "Spectrum sensing time minimizing access delay of ieee 802.11-like mac in cognitive radio networks," *IEEE Communications Letters*, vol. 15, no. 11, pp. 1249–1251, 2011.

[23] W. Yin, P. Ren, Q. Du, and Y. Wang, "Delay and throughput oriented continuous spectrum sensing schemes in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2148–2159, 2012.

[24] J. Xu, Q. Wang, K. Zeng, M. Liu, and W. Liu, "Sniffer channel assignment with imperfect monitoring for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1703–1715, 2016.

[25] M. Dabaghchian, A. Alipour-Fanid, Kai Zeng, and Q. Wang, "Online learning-based optimal primary user emulation attacks in cognitive radio networks," in *2016 IEEE Conference on Communications and Network Security (CNS)*, 2016, pp. 100–108.

[26] Q. Wang and M. Liu, "Learning in hide-and-seek," *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–14, 2015.

[27] T. L. N. Nguyen and Y. Shin, "Deterministic sensing matrices in compressive sensing: a survey," *TheScientificWorldJournal*, vol. 2013, p. 192795, 2013.

[28] A. Karimi, A. Taherpour, and D. Cabric, "Smart traffic-aware primary user emulation attack and its impact on secondary user throughput under rayleigh flat fading channel," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 66–80, 2020.

[29] D. Gözüpek, S. Buhari, and F. Alagöz, "A spectrum switching delay-aware scheduling algorithm for centralized cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1270–1280, 2013.

[30] M. Camurli and D. Gözüpek, "Channel switching cost-aware resource allocation for multi-hop cognitive radio networks with a single transceiver," in *ADHOCNETS*, 2014.

[31] M. Dabaghchian, S. Liu, A. Alipour-Fanid, K. Zeng, X. Li, and Y. Chen, "Intelligence measure of cognitive radios with learning capabilities," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[32] M. Dabaghchian, A. Alipour-Fanid, S. Liu, K. Zeng, X. Li, and Y. Chen, "Who is smarter? intelligence measure of learning-based cognitive radios," *arXiv*, 2018.

[33] A. Alipour-Fanid, M. Dabaghchian, and K. Zeng, "Self-unaware adversarial multi-armed bandits with switching costs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 6, pp. 2908–2922, 2023.

[34] IEEE Standard for Information Technology, "Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," https://www.wardriving.ch/hpneu/info/doku/802.11-1999.pdf, 1999.

[35] H. Su, Q. Wang, K. Ren, and K. Xing, "Jamming-resilient dynamic spectrum access for cognitive radio networks," in *Communications (ICC), 2011 IEEE International Conference on*, June 2011, pp. 1–5.

[36] Q. Wang, P. Xu, K. Ren, and X. y. Li, "Delay-bounded adaptive UFH-based anti-jamming wireless communication," in *INFOCOM, 2011 Proceedings IEEE*, April 2011, pp. 1413–1421.

[37] D. Roy, T. Mukherjee, M. Chatterjee, and E. Pasiliao, "Defense against pue attacks in dsa networks using gan based learning," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[38] T. R. E. T. Noam Nisan and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge University Press, 2007.

[39] J. J. Thompson, K. M. Hopkinson, and M. D. Silvius, "A test methodology for evaluating cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6311–6324, Nov 2015.

[40] A. Hess, F. Malandrino, N. J. Kaminski, T. K. Wijaya, and L. A. DaSilva, "Cognitive radio algorithms coexisting in a network: Performance and parameter sensitivity," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 4, pp. 381–396, Dec 2016.

[41] N. N. Santhosh, "Future black board using internet of things with cognitive computing: Machine learning aspects," in *2016 International Conference on Communication and Electronics Systems (ICCES)*, 2016, pp. 1–4.

[42] H. KESTELMAN, "The fundamental equation of factor analysis," *British Journal of Statistical Psychology*, vol. 5, no. 1, pp. 1–6, 1952.

[43] S. A. Mulaik, *Foundations of Factor Analysis*. Chapman and Hall/CRC, 2009.

[44] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127 – 2159, 2006.

[45] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[46] L. Li and A. Ghasemi, "IoT-enabled machine learning for an algorithmic spectrum decision process," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1911–1919, 2019.

[47] K. G. Shin, H. Kim, A. W. Min, and A. Kumar, "Cognitive radios for dynamic spectrum access: from concept to reality," *IEEE Wireless Communications*, vol. 17, no. 6, pp. 64–74, December 2010.

[48] M. J. Marcus, "Spectrum policy for radio spectrum access," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1685–1691, May 2012.

[49] J. Mitola, "Cognitive radio," Ph.D. dissertation, KTH, Teleinformatics, 2000, nR 20140805.

[50] Y. Zhao, S. Mao, J. O. Neel, and J. H. Reed, "Performance evaluation of cognitive radios: Metrics, utility functions, and methodology," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 642–659, April 2009.

[51] N. Patwari and S. K. Kasera. Crawdad utah CIR measurements. [Online]. Available: https://crawdad.org/utah/CIR/20070910/

[52] K. S. Mcgrew, "Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research," *Intelligence*, p. 10, 2009.

[53] R. J. Sternberg and S. B. Kaufman, Eds., *The Cambridge Handbook of Intelligence*. Cambridge University Press, 2011, cambridge Books Online.

[54] IBM SPSS. [Online]. Available: http://www.ibm.com/analytics/us/en/technology/spss/

[55] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The nonstochastic multiarmed bandit problem," *SIAM J. Comput.*, vol. 32, no. 1, p. 48–77, Jan. 2003.

[56] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, vol. 58, no. 5, pp. 527–535, 09 1952.

[57] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine Learning*, vol. 47, no. 2-3, pp. 235–256, May 2002.

[58] R. Arora, T. V. Marinov, and M. Mohri, "Bandits with feedback graphs and switching costs," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds. Curran Associates, Inc., 2019, pp. 10 397–10 407.

[59] O. Dekel, A. Tewari, and R. Arora, "Online bandit learning against an adaptive adversary: from regret to policy regret," in *IN PROCEEDINGS OF THE 29TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING*, 2012.

[60] N. Alon, N. Cesa-Bianchi, O. Dekel, and T. Koren, "Online learning with feedback graphs: Beyond bandits," in *Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, P. Grünwald, E. Hazan, and S. Kale, Eds., vol. 40. Paris, France: PMLR, 03–06 Jul 2015, pp. 23–35.

[61] N. Littlestone and M. Warmuth, "The weighted majority algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212 – 261, 1994.

[62] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.

[63] 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, August 2012.

[64] Y. Gai and B. Krishnamachari, "Distributed stochastic online learning policies for opportunistic spectrum access," *Signal Processing, IEEE Transactions on*, vol. 62, no. 23, pp. 6184–6193, Dec 2014.

[65] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[66] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1232–1241, April 2011.

[67] W. l. Chin, H. c. Kuo, and H. h. Chen, "Features detection assisted spectrum sensing in wireless regional area network cognitive radio systems," *IET Communications*, vol. 6, no. 8, pp. 810–818, May 2012.

[68] L. Cella and N. Cesa-Bianchi, "Stochastic bandits with delay-dependent payoffs," ser. Proceedings of Machine Learning Research, S. Chiappa and R. Calandra, Eds., vol. 108. Online: PMLR, 26–28 Aug 2020, pp. 1168–1177.

[69] C. Pike-Burke, S. Agrawal, C. Szepesvari, and S. Grünewälder, "Bandits with delayed, aggregated anonymous feedback," in *ICML*, 2018.

[70] Y. Hassan, M. El-Tarhuni, and K. Assaleh, "Learning-based spectrum sensing for cognitive radio systems," in *Journal of Computer Networks and Communications*, 2012, pp. 252–256.

[71] Y. Liao, T. Wang, K. Bian, L. Song, and Z. Han, "Decentralized dynamic spectrum access in full-duplex cognitive radio networks," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 7552–7557.

[72] A. Alipour-Fanid, M. Dabaghchian, and K. Zeng, "Self-unaware adversarial multi-armed bandits with switching costs," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 6, pp. 2908–2922, 2023.

[73] A. Alipour-Fanid, M. Dabaghchian, R. Arora, and K. Zeng, "Multiuser scheduling in centralized cognitive radio networks: A multi-armed bandit approach," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 1074–1091, 2022.

[74] J. Bonnans, *Convex and Stochastic Optimization*, ser. Universitext. Springer International Publishing, 2019.

[75] C. Daskalakis, R. Frongillo, C. Papadimitriou, G. Pierrakos, and G. Valiant, *Algorithmic Game Theory Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2010.

[76] S. P. R. Susan E. Embretson, *Item Response Theory for Psychologists*. LEA, 2000.

[77] SAS Institute Inc., *SAS/STAT Software, Version 9.1*, Cary, NC, 2003. [Online]. Available: http://www.sas.com/

[78] R. Fourer, D. M. Gay, and B. W. Kernighan, "A modeling language for mathematical programming," *Management Science*, vol. 36, no. 5, pp. 519–554, 1990.

[79] M. C. Grant and S. P. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, V. D. Blondel, S. P. Boyd, and H. Kimura, Eds. London: Springer London, 2008, pp. 95–110.

[80] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.