Combating Jammers in Cognitive Radio Networks

By
Khalid Ibrahim
102-FET/Ph.D.EE/F15

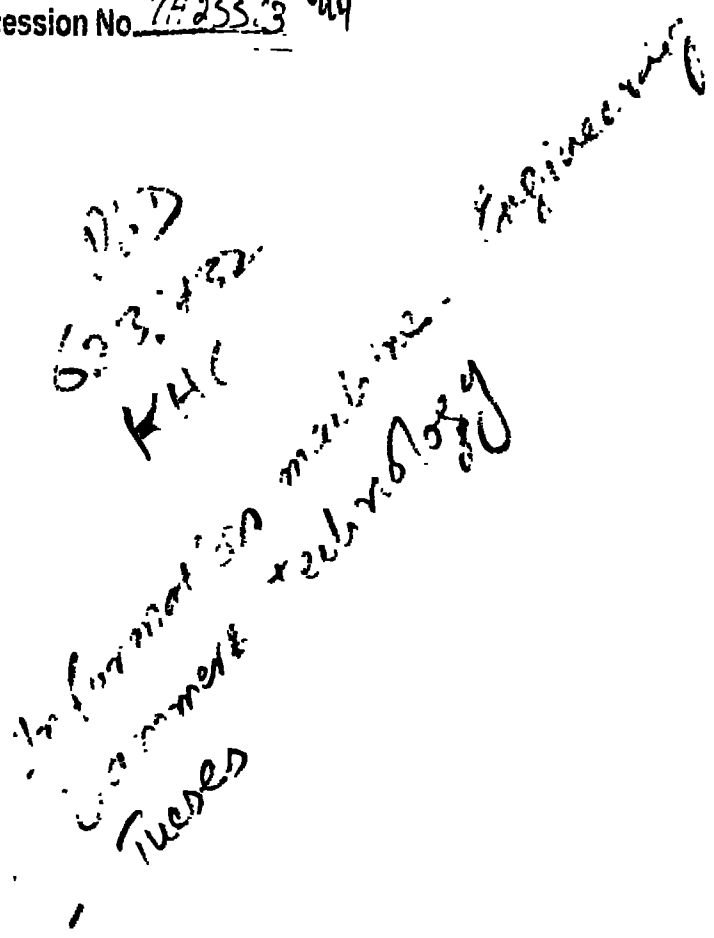
Supervisor
Dr. Aqdas Naveed Malik

A dissertation submitted to I.I.U. in partial fulfilment of the requirements for the degree of

Philosophy of Doctorate

Department of Electrical Engineering,
Faculty of Engineering and Technology (FET),
International Islamic University
Islamabad, Pakistan
2022

Accession No. 74255.3 44



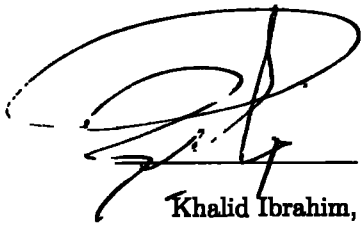
Copyright Notice

Declaration

I, *Khalid Ibrahim* declare that this thesis titled "Combating Jammers in Cognitive Radio Networks" and the work presented in it are my own and have been generated by me as a result of my original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Ph.D. degree at IIUI
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at IIUI or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my work
5. I have acknowledged all main sources of help
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself



Khalid Ibrahim,
102-FET/Ph.D.EE/F15

CERTIFICATE OF APPROVAL

Title of Thesis: Combatting Jammers in Cognitive Radio Networks

Name of Student: KHALID IBRAHIM

Registration No: 102-FET/PHDEE/F15

Accepted by the Department of Electrical & Computer Engineering, Faculty of Engineering and Technology, International Islamic University (IIU), Islamabad, in partial fulfillment of the requirements for the Doctor of Philosophy degree in Electronic Engineering.

Viva voce committee:

Prof. Dr. Aqdas Naveed Malik (Supervisor)
Professor, DEE, FET, IIU Islamabad.

Dr. Muhammad Amir (Internal)
Professor DECE, FET, IIU Islamabad.

Dr. Noman A. Khan (External-I)
Chairman DECE, CASE, Islamabad.

Dr. Rab Nawaz (External-II)
Project Director, NESCOM, Islamabad.

Dr. Shahid Ikram (Chairman, DECE)
Associate Professor DECE, FET, IIU Islamabad.

Prof. Dr. Nadeem Ahmad Sheikh (Dean, FET)
Professor DME, FET, IIU Islamabad.

Dedication

This thesis is dedicated to my beloved mentor, renowned researcher and source of motivation

Prof. Dr. Ijaz Mansoor Qureshi

(May his soul rest in peace.)

Awards and Grants

1. Award of scholarship by Higher Education Commission (HEC) Pakistan under "International Research Support Initiative Program (IRSIP)" for six months research fellowship at Next Generation Wireless (NGW) lab of the University of Southampton, UK covering:
   - Stipend (4500 UK Ponds)
   - Bench fee waived off.
   - Travel grant (15,0000 PKR)
2. Full fee waiver for Ph.D. studies by the International Islamic University Islamabad (IIUI), Pakistan.

Acknowledgements

In the name of Allah (Subhanahu Wa Ta'ala), who is the most gracious and the merciful. I would like to thank Allah for giving me strength and patience to complete this research work. Peace and blessings of Allah be upon His last Prophet Muhammad (Sallulah-o-Alaihihe-Wassalam) and all his Sahaba (Razi-Allah-o-Anhu) who dedicated their lives for Dawah and spread of Knowledge.

I am truly grateful to my supervisor Prof. Dr. Aqdas Naveed Malik, who has supported me throughout this research project. I'd like to express my thanks to my co-supervisor, Prof. Dr. Ijaz Masorr Qureshi (Late), whose inspiration, ideas and efforts make it possible to complete my higher studies. He has been a role model for me and many others in teaching, research and other aspects of life.

I offer sincere thanks to my seniors and colleagues Dr. Muhammad Naeem, Dr. Zeeshan Kaleem, Dr. Umer Chughtai, Dr. Ayaz Ahmed, Dr. Maaz Rehan, Dr. Mubashir Hussain Rehmani, Dr. Abdullah Waqas, Dr. Sadiq Ahmed, Dr. Athar Waseem and Dr. Sheroz Khan for their never-ending support and fruitful and healthy

research discussions. I am also indebted to all anonymous reviewers and researchers in all revision of the published papers and thesis for their valuable comments and suggestions to improve the quality of research work.

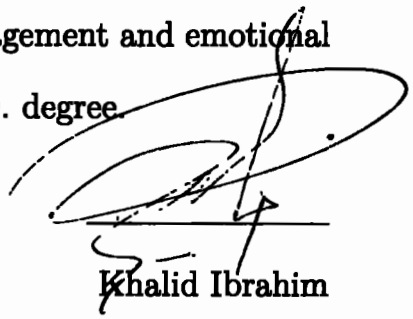
I would like to pay special thanks to Dr. Soon Xin Ng (Micheal) from the University of Southampton as foreign supervisor during my visit to Next Generation Wireless Lab, UoS, UK. I would like to acknowledge the support of Dr. Hassan Mehmood from Quaid-i-Azam University, Islamabad, for his fruitful discussion during the completion of the thesis.

I am thankful to my friends Engr. Belawal Behram, Engr. Abdullah Shoukat, Dr. Abdul Rehman Al-Salehi, Engr. Fahad Munir, Engr. Ahmed Saleem, Muhammed Ibrahim, Ahmed Saeed, Zunnorain Amer Ali Baig, Dr. Zeeshan Aslam, Dr. Khizar Mehmood, and Dr. Muhammad Muzammil for their help, support, and motivation.

I would like to acknowledge the support of Higher Education Commission for funding my visit to Next Generation Wireless Lab, University of Southampton as International Research Support Initiative Program (IRSIP) scholar. I would like to thank all my friends Mehmet Abdul Rahim Sen (Turkey), Dr. Muhammad Jamal (Egypt), Dr Muhammad Shoib (Pakistan), Dr. Mansoor Ahmed (Egypt), Hisham Ghabra (Syria), Dr. Saif Zyad (Aljeria), and many others during my visit in Southampton, specially the ISOC team for

providing me a good company in the absence of my family members. I am also thankful to International Islamic University Islamabad, Pakistan for providing me a full fee waiver during the Ph.D. studies. I am thankful to administration at department, and the university level for their kind support.

I will always be indebted and thankful to my parents, brothers and sister for their love and support throughout my life. And special thanks to my kids Saleema, Haleema and Anas for their smiling faces, unbridled laughter and their excitement towards 'baba' made me stronger day by day. Last but not the least pay my gratitude to my wife, Khadeejah, for her patience, encouragement and emotional support during every single stage of my Ph.D. degree.



Khalid Ibrahim

Abstract

Cognitive Radio (CR) provides a promising solution to the spectrum scarcity problem in dense wireless networks, where the sensing ability of cognitive users helps acquire knowledge of the environment. However, cognitive users are vulnerable to different types of attacks, due to its shared medium. In particular, jamming is considered as one of the most challenging security threats in CR networks. In jamming, an attacker jams the communication by transmitting a high-power noise signal in the vicinity of the targeted node. The jammer could be an intelligent entity capable of exploiting the dynamics of the environment. This work presents a machine-learning-based anti-jamming technique for CR networks to avoid a hostile jammer, where both the jamming and anti-jamming processes are formulated based on the Markov game framework. In the proposed framework, secondary users avoid the jammer by maximizing its payoff function using an online, model-free reinforcement learning technique called Q-learning. A realistic mathematical model is proposed, where the channel conditions are time-varying and differ from one sub-channel

to another, as in practical scenarios.

Anti-jamming in cognitive radio networks is mainly accomplished using machine learning techniques in frequency, code, power and rate domains. With the improvement in communication technologies, the capabilities of adversaries are increased as well. The intelligent jammer knows the rate at which users are transmitting data and is based on the attractiveness factor of each user. The higher the data rate of a secondary user, the higher its attractiveness to the rate-aware jammer. In the second part of this work, a dummy user is introduced in the network as a honeypot for jammer to attract the attention of the jammer. Moreover, a novel game-theoretic anti-jamming deception method based on rate adjustments is presented to increase the bandwidth efficiency of the whole cognitive radio-based communication system. A defensive anti-jamming deception mechanism is devised to decoy the attacker to protect the rest of the network from the impact of the attacker. The simulation results show a significant improvement in performance using the proposed solution.

List of Publications

Published Papers

1. **K. Ibrahim**, S. X. Ng, I. M. Qureshi, A. N. Malik and S. Muhaidat, 'Anti-Jamming Game to Combat Intelligent Jamming for Cognitive Radio Networks,' *IEEE Access*, vol. 9, pp. 137941-137956, 2021, doi: 10.1109/ACCESS.2021.3117563.

2. **K. Ibrahim**, I. M. Qureshi, A. N. Malik, and S. X. Ng, "Bandwidth-efficient frequency hopping based anti-jamming game for cognitive radio assisted wireless sensor networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–5.

3. **K. Ibrahim**, I. M. Qureshi, A. N. Malik and A. Waseem, 'COVER FIRE: Deceptive Anti-Jamming using Pseudo Secondary Users in Cognitive Radio Networks,' 2021 International Conference on Engineering and Emerging Technologies (ICEET), 2021, pp. 1-4, doi: 10.1109/ICEET53442.2021.9659646.

4. **K. Ibrahim**, A. M. Alnajim, A. N. Malik, A. Waseem, S. Alyahya, M. Islam, S. Khan, "ENTICE TO TRAP: Enhanced Protection Against a Rate-Aware Intelligent Jammer in Cognitive Radio Networks," *Sustainability*, vol. 14, no. 5, p.2957,2022.

This Ph.D. thesis is based on the above mentioned research papers.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AWGN	Additive White Gaussian Noise
CH	Channel Hopping
CR	Cognitive Radio
CRN	Cognitive Radio Network
CRSN	Cognitive Radio Sensor Networks
DoS	Denial of Service
DSSS	Direct Sequence Spread Spectrum
FCC	Federal Communication Commission
FH	Frequency Hopping
FHSS	Frequency Hopping Spread Spectrum
GSG	General Sum Game
HLA	Hierarchical Learning Algorithm
IL	Independent Learner
JAL	Joint Action Learner

KTH	Kungliga Tekniska Högskolan (Royal Institute of Technology)
MDP	Markov Decision Process
MARL	Multi-Agent Reinforcement Learning
MLE	Maximum Likelihood Estimation
NE	Nash Equilibrium
OLOFS	One Leader One Follower Stackelberg Game
PC	Power Control
PSD	Power Spectral Density
PSU	Pseudo Secondary User
PU	Primary User
PUEA	Primary User Emulation Attack
RA	Rate Adaptation
RF	Radio Frequency
SDR	Software Defined Radio
SSDFA	Spectrum Sensing Data Falsification Attack
SU	Secondary User
SG	Stochastic Game

SLMFS	Single Leader Multiple Follower Stochastic Game
SE	Stackelberg Equilibrium
SNR	Signal to Noise Ratio
WSN	Wireless Sensor Network
VNET	Virtual Network

List of Notations

α	Learning rate
β	Regret factor of jammer
η	Bandwidth efficiency (bps/Hz)
γ	Discount factor
a_l^n	Action of the n^{th} secondary user at l^{th} channel
c_j^m	Action of the m^{th} jammer at l^{th} channel
$C_{l,t}(a^n, c_j^m)$	Channel capacity of the n^{th} secondary user at l^{th} channel at time slot t
$\mathcal{E}_l^n$	Transmission cost of the n^{th} secondary user at l^{th} channel
$\mathcal{E}_{jl}^m$	Transmission cost of the m^{th} jammer at l^{th} channel
$\mathcal{T}_l^n$	Transmission gain of the n^{th} SU at the l^{th} channel
$\mathcal{T}_{jl}^m$	Jamming gain of the m^{th} jammer at the l^{th} channel

$\mathcal{J}_l^n$ Transmission loss of the n^{th} secondary user at the l^{th} channel due to jamming

l Index of the sub-channels

$\mathcal{L}$ No. of sub-channels in the network

$\mathcal{H}$ No. of primary users in the network

$\mathcal{M}$ No. of jammers in the network

$\mathcal{N}$ No. of secondary users in the network

$Q(s^k, a_l^k)$ Q table entries for the secondary user at the l^{th} channel in the k^{th} time slot

$Q^j(s^k, c_j^k)$ Q table entries for jammer at the l^{th} channel in the k^{th} time slot

$\mathcal{A}(s)$ Action space of secondary users

S State space of the game

$\mathcal{U}_{l,t}^n(a^n, c_j^m)$ Utility of the the n^{th} secondary user at l^{th} channel at time slot t

$\mathcal{U}_{jl,t}(a^n, c_j^m)$ Utility of the jammer at the l^{th} channel at time slot t

$x_{l,t}(a^n, c_j^m)$ Switching function of the n^{th} secondary user at l^{th} channel at time slot t

$\mathcal{P}_{l,t}^n$ Average received power at the l^{th} channel by the n^{th} secondary user at time slot t

CHAPTER 1

Introduction

1.1 Significance of Wireless Communication

Wireless communication has grown very fast in the last decade. Consequently, wireless communication systems have become inevitably related to several applications and a range of devices e.g., smart phones, laptops, and IPADs. In addition, new wireless applications like wireless sensor networks (WSN), vehicular ad-hoc networks (VANETs), smart home appliances, smart grids, remote telemedicine, and numerous others are materializing from research thoughts to tangible systems. Moreover, wireless communications is advantageous due to accessibility, easy installation, wider reach, flexibility, efficiency and cost effectiveness. Wireless communication is the most absolute and vital requirement of this present era, and it is the staircase for further advancements in the field of digital communication. However, with the implausible growth in the systems and applications, the availability of the wireless spectrum, the natural source

which provides all this communication, is limited. Wireless communications have enabled billions of individuals to connect to the Internet and benefit from today's digital economy.

1.2 Overview of Cognitive Radio Networks

This enormous growth in applications is inevitable now, and limitation is evident from the frequency allocation charts for United States and frequency allocation charts for Pakistan. Nevertheless, statistics taken out by the Federal Communication Commission with the help of experiments in various countries show that most radio frequency bands are not used the majority of the time or are underutilized. So there is a need to address the problem which could solve the underutilized or inefficient use of spectrum, i.e., how and when to use it [3], [4].

As this inadequate natural resource will ultimately get halted for the users coming in the near future, a solution was required for the spectrum scarcity. The solution was proposed by Joseph Mitola from KTH in his Doctoral Dissertation in 1999 [5], [6]. This dissertation presents a conceptual overview of cognitive radio as an exciting multidisciplinary subject. Federal Communications Commission (FCC) defines cognitive radio as "A system or radio that sniffs its functioning Electromagnetic atmosphere and capable of vigorously

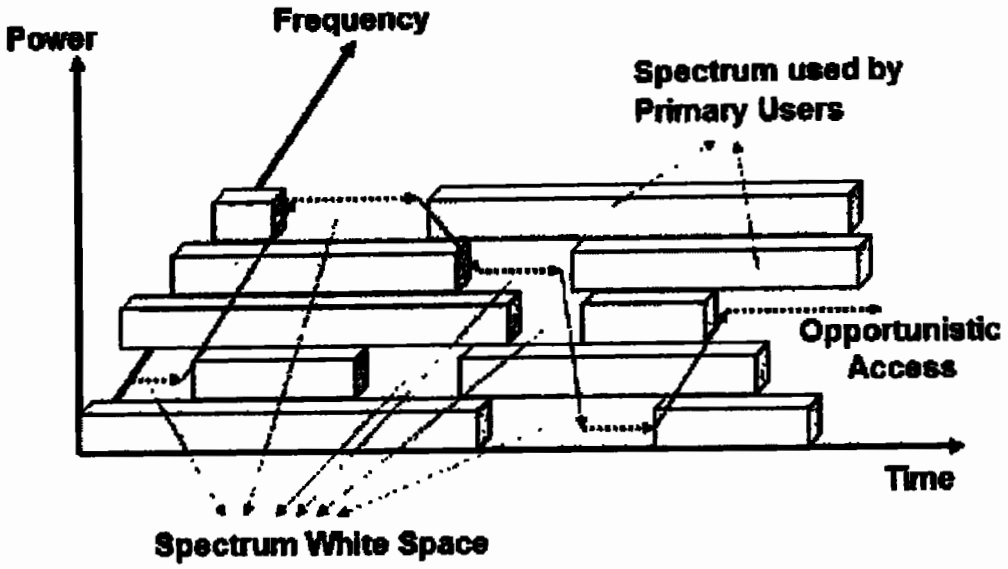


Figure 1.1: Opportunistic Access of Spectrum in CRN. The empty white (unshaded region) shows the white space available for CR users while PU is active in the shaded region [6].

and separately regulate its radio working parameters to adapt intrusion, smooth the progress of interoperability, and access secondary users." Cognitive radio is aware of its surrounding RF environment learns, reason, decide and adapt to the external conditions [7]. The aim is to utilize the spectrum efficiently and carry out reliable and uninterrupted wireless communication.

CR is practically implemented on the hardware referred to as Software Defined Radio (SDR). Analysis of the Radio Scene, Recognition and Identification of channel and Power Control of Transmission are three important roles of the cognitive cycle [8]. The learning and reasoning abilities of cognitive radios are embedded in the cognitive engine, the core of a CR, as depicted in Figure 1.2. The task of the cognitive engine is to coordinate the actions of CR

using machine learning algorithms [9]. A CR node can switch its operational frequency to the dynamic RF environment. CR nodes can access spectrum white spaces that are not being utilized by primary users, as depicted in Figure 1.1. Two types of users in Cognitive Radio Networks are Primary User (PU) and Secondary User (SU). When the primary user is not available, the cognitive user can use its resources, i.e., the un-utilized spectrum. Competition takes place among secondary users to access the available spectrum for primary users. The absence of a primary user in certain time and space results in the unoccupied frequency termed as White space [10]. Wireless microphones and TV towers contribute to the most significant proportion of white space [11]. Federal Communication Commission (FCC) in the USA has adapted the rule for unlicensed use of Television White spaces due to its underutilized spectrum [3]. Spectrum sensing plays an essential role for dynamic spectrum access in CRs [12]. The task of spectrum sensing comprises two significant factors. One is to protect the rights of the licensed user, for this objective CR continuously observes the frequency environment and identifies the spectrum holes, that is, it enables the SUs to exploit the unused PU spectrum.

It is very important to understand the activity of primary user so that SU can make a decision accordingly. The performance of cognitive radio networks is highly dependent upon the activity of

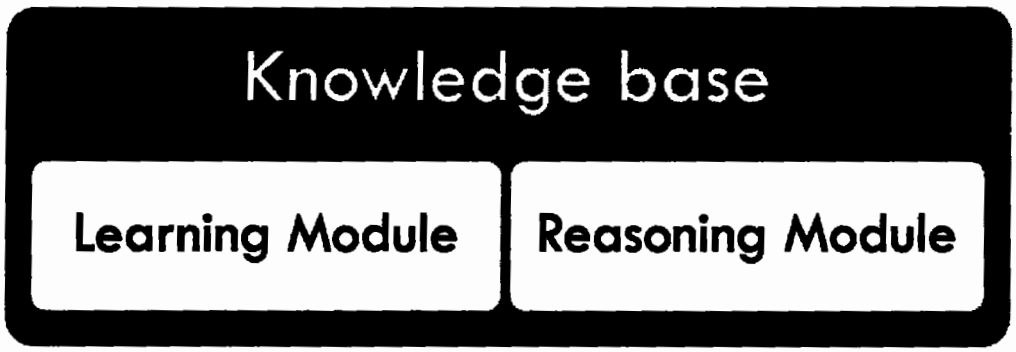


Figure 1.2: Cognitive Engine, the heart of cognitive radios, comprises three interactive modules knowledge base, learning module and reasoning module.

primary radio users [13]. In literature, PR user activity is modeled based on Markov process, queuing theory, time series, and ON/OFF periods. Other models are the Bayesian model, Event-based random walk model, PR user presence probability from historical statistics, Partial periodic pattern mining (PPPM), First-difference filter clustering. There are three paradigms of cognitive radio networks, e.g., Underlay, Overlay and Interweave paradigms.

1.3 Security Issues in Cognitive Radio Networks

Cognitive radio networks are incredibly vulnerable to malicious attacks, partly because secondary users do not own the spectrum, and hence their opportunistic access cannot be protected from adversaries. Moreover, highly dynamic spectrum availability and often distributed network structures make it challenging to implement adequate security countermeasures [14]. In addition, as cognitive radio networks benefit from technology evolution to be capable of utilizing

spectrum adaptively and intelligently, the same technologies can also be exploited by malicious attackers to launch more complicated and unpredictable attacks with more significant damage [15, 16]. Therefore, ensuring security is paramount to the successful deployment of cognitive radio networks. More explicitly, jamming attacks, Denial of Service (DoS) attacks [17, 18], Primary User Emulation Attack (PUEA) [19–21], Spectrum Sensing Data Falsification Attack (SSDFA) [22], exploitation of common control channel security [23] and collaborative jamming [24] are well-known attacks in cognitive radio networks. However, the major concern of this research work is to combat jamming attacks in cognitive radio networks.

1.4 Introduction to Game Theory

Game theory is widely used in the literature to model the competitive environment between jammer and secondary user. A game is a mathematical model for interactive situation where players have to make decisions based on the payoffs. It provides the formal framework that helps generate useful information for analysis purposes. It is the situation in which players make strategic decisions that take into account each other actions and responses [25].

Figure 1.3: Key elements of a non-cooperative frequency hopping game, aiming is to reach Nash equilibrium.

1.4.1 Key Elements

The key elements of the game are players, strategies, payoffs, information and rationality as depicted in Figure 1.3. These terms are defined below.

Players: Those who are interacting. In this case, two players are SU and jammer.

Strategies: Rules or plan of action of each player for playing game hop, stay.

Payoffs: What are the players gaining after adapting certain strategies? And the optimal strategy is the one that maximizes the player's payoff.

Information: What do the players know? Completer information is in which each player know every aspect of the game while in perfect information, the player only knows the previous actions taken by all other players.

Rationality: Players are assumed rational to take the best alternative in the set of possible options. It helps narrow down the possible

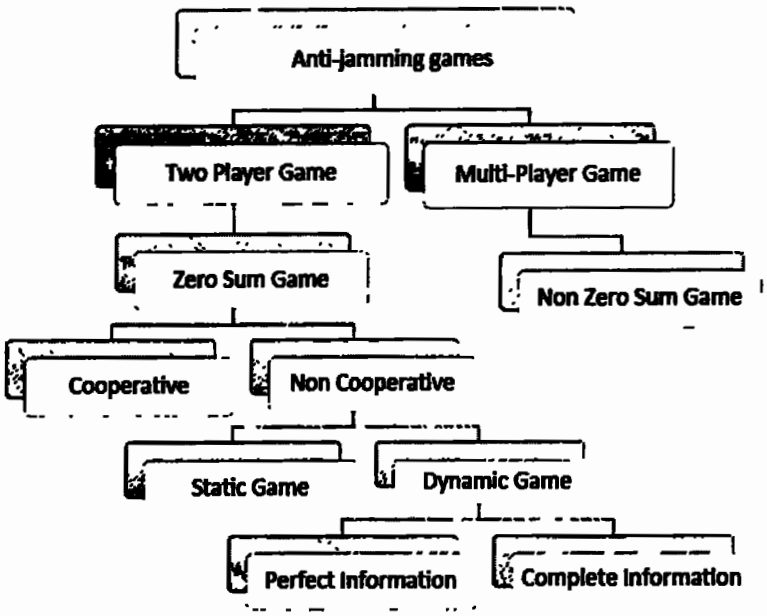


Figure 1.4: The game-theoretical anti-jamming tree describes the zero-sum non cooperative stochastic game between SU and jammer.

decisions.

Nash Equilibrium: It is an action vector from which no player can profitably unilaterally deviate. N players interact to take a set of actions A_i for each player. Each action has a certain outcome described as utility function U_i. An action A_i is Nash Equilibrium if for every i player out of N players satisfy the following inequality

$$U_i(a_i, a_(-i)) > U_i(b_i, a_(-i)) \quad \forall b_i A_i \tag{1.4.1}$$

Players exhibit rational behavior and adapt the strategy to maximize their payoffs. Rational choice theory is an economic principle which states that individuals always make prudent and logical decisions. These decisions provide people with the greatest benefit or satisfaction, given the choices available, and are also in their highest self-interest. Game theory reduces the complexity of adap-

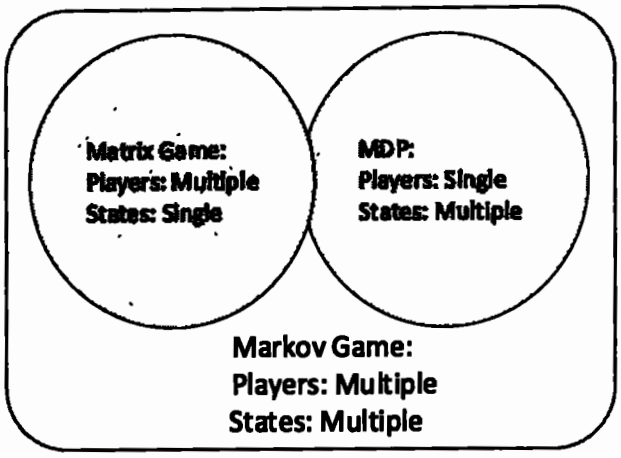


Figure 1.5: Relationship between MDP, Markov games and matrix games.

tation algorithms in large cognitive networks [9]. Figure 1.4 shown a game-theoretical anti-jamming tree describing the zero-sum non cooperative stochastic game between SU and jammer.

1.4.2 Relationship Between Stochastic Game and Markov Decision Process

A Stochastic Game (SG) is the natural extension and generalization of the Markov Decision Process (MDP) to multi-agent systems [9, 26].[1] The relationship between Markov games, matrix games and MDP is shown in Figure 1.5. SG provides a framework for multi-agents in multi-agent reinforcement learning (MARL). In this contribution, a stochastic anti-jamming game is developed between two players of conflicting interests. [See Appendix A for the details about game theory.]

[1]Stochastic games are also called Markov games, which essentially are n-agent MDP [27].

1.5 Reinforcement Learning

The Q-learning is a value-based reinforcement learning algorithm, which uses a Q table to maximize the utility [28]. It is a mechanism learn the sub-channel selection strategy effectively. The Q function $Q(s^k, a^k)$ at stage k, is the expected discounted payoff when the SU takes the action a^k at the state s^k. More specifically, the Q value is the estimation of the expected sum of the discounted payoff [29]. Hence, an SU can consider the Q value in a bimatrix game at stage k as the expected sum of the discounted payoffs given in Equation (4.1.3). Given the Q function $Q(s^k, a^k)$, the SU can find the value of the game from:

$$V(s^k) = \max_{a^k} Q(s^k, a^k). \tag{1.5.1}$$

After an action a^k is taken, the SU would receive an immediate payoff $R(s^k, a^k)$, which is then used to update the Q table. Specifically, the Q function can be approximated as:

$$Q(s^k, a^k) = R(s^k, a^k) + \sum_{s^{k+1} \in S} Pr(s^{k+1}, a^{k+1} | s^k, a^k) V(s^{k+1}), \tag{1.5.2}$$

where $Pr(s^{k+1}, a^{k+1} | s^k, a^k)$ is the transition probability from state s^k to s^{k+1}. Q-learning is a model free learning algorithm adapted to learn the optimal policy without explicitly knowing the model. The intuition behind Q-learning is to approximate the unknown transition probability in Equation (1.5.2) by the empirical distribution of

states that have been visited as the game proceed [30].

1.5.1 Multi-Agent Reinforcement Learning (MARL)

The study of how numerous agents interact in a common environment is known as multi-agent reinforcement learning (MARL). MARL is a sub-field of reinforcement lea ring and is becoming popular. When these agents engage with the environment and each other, they can collaborate, coordinate, compete, or learn to complete a job collectively.

1.5.2 Types of Learners in MARL System

A multi-agent reinforcement learning (MARL) algorithm is the independent learning (IL) algorithm where the learner can take action individually and do not consider the actions taken by other agents. There are two types of learners in the MARL setting, namely the independent learner (IL) and the joint action learner (JAL) [31]. IL uses Q-learning in a classical setting, ignoring the other agents. More specifically, it assumes that the other agents are part of the environment. A MARL algorithm is an IL algorithm if the learner would take action individually and would not consider the actions taken by other agents. The IL algorithm is an appropriate learning method if the agent is unaware of the other agents in the system and their actions [32]. If the above condition does not hold, an agent can still

ignore the presence of other agents to justify the application of the IL algorithm. Learning is relatively simple for IL, as it only learns its actions [31].

The total number of entries that an IL agent needs to learn is given by $m\times \mid A^1 \mid$ for an n-agent system, which has m states of the game, where $\mid A^1 \mid$ is the size of the action space of player 1. Since the IL ignores the actions of other agents, the complexity of the IL agent is linear as given by $m\times \mid A \mid$.

A JAL is an agent that learns the environment in the presence of other agents and then updates its Q values based on the joint actions of all the agents in the system. This infers that the agent knows the rewards of all other agents, and its experience is of the form $< \mathbf{a}, \mathbf{r} >$ where $\mathbf{a} = a_1 \times a_2 \times ... \times a_n$ is the joint action of all the agents and $\mathbf{r}$ is the reward of the joint action $\mathbf{a}$. The complexity of the JAL is exponential, as it has to learn all possible actions of all the agents in the system. A toy example of a bi-matrix zero-sum game is represented in Table 1.1. If player A is a JAL it has to learn all joint actions, i.e., $(a_0, b_0), (a_0, b_1), (a_1, b_0), (a_1, b_1)$, while if player A is an IL, then it needs to learn its own actions only, i.e., a_0 and a_1.

Table 1.1: Toy example of a game for the comparison of IL and JAL

		Player B	
		b_0	b_1
Player A	a_0	5	0
	a_1	0	5

In [26], Littman considers a JAL in the minmax settings. The states S, action set A, and the opponent action set O are the variables from the environment. In [27], Hu and Wellman give a more general form of a JAL for the general sum game (GSG) in which each JAL player assumes those other players are rational and that other agents will take actions according to their own believes about other agents. More explicitly, each agent maintains a belief about all other players in the network and this belief is maintained in the form of a separate Q table, which results in increased complexity for the learning system. The total number of entries that a JAL agent needs to learn is given by $m \times \mid A^1 \mid \times \mid A^2 \mid$ for a two-agent system having m states of the game, where $\mid A^1 \mid$ is the size of the action space of player 1. Assuming equal action space for each of the n-agents, the exponential complexity of the JAL agent is $m \times \mid A \mid^n$. In fact, even though JALs have much more information at their disposal, they do not perform much different from ILs in the straightforward application of Q-learning to multi-agent systems [31]. Both the SU and jammer players are taken as ILs, where each IL would apply the Q-learning algorithm in the classical setting while ignoring the action of the other agent. More specifically, each agent assumes that the other agents are part of the environment. IL algorithm is an appropriate method of learning if the agent is unaware of the other agent in the system, hence do not know the actions of other agents.

Learning is relatively simple for IL, as it only has to learn its actions [31]. A multi-agent reinforcement learning algorithm (MARL) is the IL algorithm where the learner can take action individually and do not consider the actions taken by other agents [32].

Theorem 1[33]: An IL agent in a MARL setting, following the Q-learning update rule, will converge to the optimal Q-function with unit probability.

1.6 Motivation

CRN is a promising technology to cope with the scarcity issue of the electromagnetic spectrum, which is a natural resource. Traditional wireless radio communication works on fixed frequency slots, resulting in overcrowding in certain portions of the electromagnetic spectrum while other portions are underutilized. CR is aware of its surrounding Radio Frequency (RF) environment. It learns, reasons, decides, and adapts to external conditions to efficiently utilizing the radio spectrum and carry out reliable and uninterrupted wireless communication [7, 8]. Furthermore, CR could provide opportunistic access to spectrum holes to solve the intermittent use of radio spectrum using machine learning algorithms [9, 34–37]. Therefore, ensuring security is of paramount importance to the successful deployment of cognitive radio networks.

In traditional wireless communication systems, Frequency Hoping Spread Spectrum (FHSS) and Direct Sequence Spread Spectrum (DSSS) are widely used to thwart jammers [18, 38]. Due to dynamic spectrum mobility [39], it is impossible to directly apply these techniques in cognitive radio technology to combat the hostile jammers. Since jamming attacks drastically degrade the performance of cognitive radio, some effective mechanism is required to mitigate the effects of jammer by avoiding and deceiving the jammer.

1.6.1 Jamming Attack

Jamming is a major threat to CRN [40–43]. Jammers disrupt wireless communication by generating high-power noise across the entire bandwidth near the transmitting receiving nodes. As a result, the communication channels either cannot be accessed or the signal-to-noise ratio (SNR) in these channels heavily deteriorates. The jammer model in a CRN is shown in Figure 1.6, where the jammer disrupts the wireless communication by generating high-power noise, causing narrow-band interference on a single sub-channel at a time near the transmitting and receiving nodes [44]. Intensive jamming could result in either total disruption of the wireless communication or a very low SNR that does not allow secondary users (SUs) to communicate successfully. Since jamming attacks drastically degrade the performance of cognitive radio, some effective mechanisms are required to

detect their presence and avoid them. More specifically, this research focuses on anti-jamming techniques using frequency hopping (FH). The interested readers may refer to [45, 46] for details concerning jamming detection.

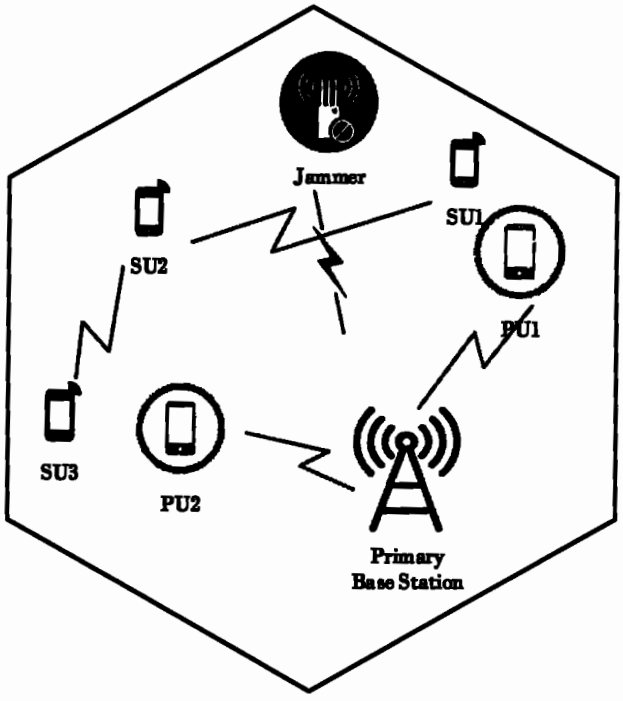


Figure 1.6: The model of the jammer in a cognitive radio network

1.6.2 Intelligent Jamming Attacks

Existing naive jammers mostly rely on high power and frequent transmission of jamming signals which is not practical for power constraint jammers. Moreover, this kind of high power and frequent jammers also ease jamming detection. The adaptability of CR could provide intelligence to spectrum sensing and spectrum decision [36].

On the other hand, the adversary can also maneuver the same features intelligently to create more harm to the underlying CR Network (CRN) [29, 47, 48]. A more powerful intelligent jammer is considered which targets the users based on the attraction factor of each user. The attraction factor is proportional to the rate at which the communication is carried out. Hence targeting the highest impact communications in the cognitive radio network [49]. Here are the other reasons to consider rate-aware jammers.

- It is easier for the rate-aware intelligent jammer to target a few symbols of higher data rates resulting in very efficient selective jamming.
- A targeted, efficient attack will force the secondary user to communicate at a lower rate by jamming all communication at higher rates as shown in [49].
- The low data rates will result in network saturation, which causes higher collision probability [49].

1.6.3 Frequency Hopping

The most widely considered method for reducing the effects of jamming attacks is frequency hopping. Frequency hopping is a technique for quickly switching between many frequency channels when sending radio communications. The ease of implementation and robustness

against interference and jamming attacks have made proactive frequency hopping more popular. Frequency hopping is particularly effective when the number of orthogonal channels supported is substantially larger [50].

1.6.4 Rate Adaptation

Rate adaptation schemes in the litterateur usually adjust the physical layer transmission rate according to the channel conditions, ideally choosing high data rate by adapting more robust modulation and coding schemes (MCS) for good SNR channels and low transmission rates for poor channel conditions [51, 52]. Transmitting at the modulation scheme with a higher data rate will increase the probability of getting jammed due to rate-aware jammers in the network. On the other hand, transmitting at low rates increases the robustness and reliability against jamming but will reduce the throughput of the system [53]. Therefore, an adequate data transmission rate is required for effective transmission while avoiding the jammer [49, 54].

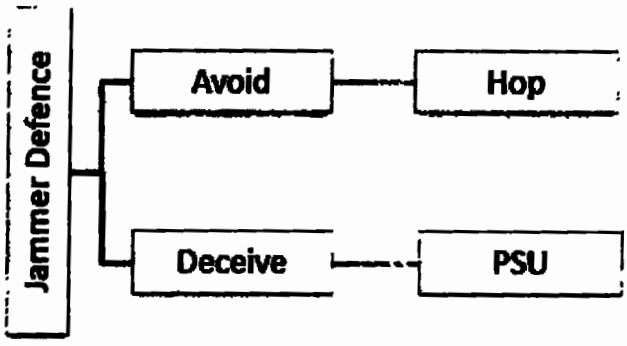


Figure 1.7: Anti-jamming games are widely classified into traditional and deception-based games.

1.7 Problem Statement

Problem statement of this research is divided into two parts. The first part is to devise a game-theoretic an anti-jamming scheme against intelligent jammer. The second part is extended further against a more hostile rate-aware intelligent jammer. The details are given below.

1.7.1 Problem-I: Adversarial Anti-Jamming Game

The above discussion shows that the research community has contributed much research towards anti-jamming for CRN in the frequency domain. However, most literature has assumed a fixed strategy for the jammer, which is not changing with time. With the development and technological advancement in the cognitive radio networks, it is highly conceivable that a jammer will also manoeuvre

its attacking strategies intelligently. Hence, there is a need for an intelligent anti-jamming strategy for CRN. An intelligent jammer is cognitive in nature, having the ability to learn, reason, and adjust its strategies against SU for maximum damage to the CRN.

1.7.2 Problem-II: Deception-based Anti-Jamming Game

Deception empowers network administrators by thoroughly defending against attacks from both external parties and hostile insiders, properly warning when something is wrong, and offering precise threat intelligence for quick remediation. When an intrusion is detected, they can observe how the intruder moves around the infrastructure and what resources they appear to be targeting. They can then investigate specific network parameter which was targeted to deceive the attacker. Therefore, there is a need for a deception strategy against an intelligent jammer to waste its resources hence protecting the overall network.

1.8 Objectives of Research

The jammer is cognitive, it also looks for white space. There is a competition for spectrum occupancy between the secondary user and cognitive jammer. The two objectives of the research are listed below.

1. The objective is to devise an optimal hopping scheme to pick the optimal channel before the jammer catches up to it.

2. The second objective of the research is to develop a deceiving mechanism for jammers, so that the jammer wastes its energies while cognitive users enjoy its transmission thereby decreasing the probability of being jammed.

The two objectives of the research are shown in Figure 1.7. The following performance criterion will be used to evaluate the performance of the proposed anti-jamming scheme.

- Jamming Probability under cognitive attacks.
- The average payoff of different strategies

Game theoretic analysis is used to devise an anti-jamming mechanism to combat against a variety of jamming levels. Game theory is widely used in wireless communications to solve communication problems like resource allocation [55–57], packet relaying [36], and anti-jamming communication [29, 48, 58–60] , as shown in Figure 1.8.

In traditional wireless communication systems, Frequency Hoping Spread Spectrum (FHSS) and Direct Sequence Spread Spectrum (DSSS) are widely used to thwart jammers. Due to dynamic spectrum mobility, it is impossible to directly apply these techniques

in cognitive radio technology to combat hostile jammers. Therefore, this research focuses on machine learning based algorithms to provide safety of the network in the presence of hostile jammers. Moreover, Deception in cyber security of wireless communication is largely adapted for the following three reasons. (1) For the detection of the attacker, (2) for information about the intelligence of the attacker, (3) for confusing the adversarial user to waste its resources on the sweetener. Since detection is not the focus of this article, and the intelligence of the jammer is assumed to be a posteriori knowledge of the user, therefore, the focus of this research work is to confuse the attacker between a legitimate target and deceptive sweetener by using a deception strategy in the CRN. The scope of the thesis is limited to machine learning-based anti-jamming techniques to combat and deceive the jammer, hence providing enhanced protection against intelligent jammers in cognitive radio networks.

1.9 Organization

The rest of the thesis is organized as follows. Chapter 2 provides a detailed and state-of-the-art literature survey on anti-jamming techniques in CRN. Moreover, this chapter provides a review of the previous research on deception-based defense strategies. It also provides the comparison table as evidence of the novelty presented in this research work. Chapter 3 explains the system and adversary

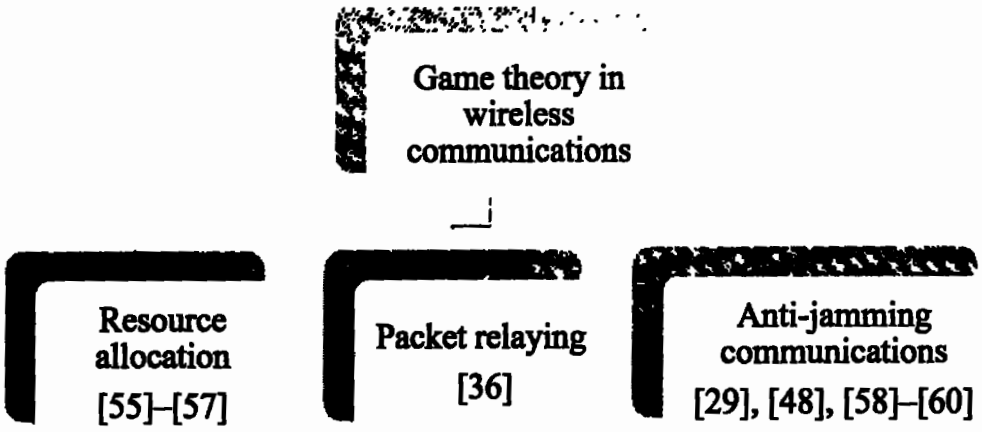


Figure 1.8: Taxonomy of game-theoretic applications in wireless communications.

models in detail. The system and adversarial model are divided into two parts in this chapter, part-I and part-II. Chapter 4 provides an anti-jamming game formulation against a random and an intelligent jammers, respectively. Moreover results and discussions are provided based on the simulation results. Chapter 5 provides a deception based anti-jamming mechanism to combat intelligent jammers. Moreover results and discussions are provided based on the simulation results. Finally, chapter 6 concludes the thesis with the future works in the light of limitations.

CHAPTER 2

Literature Review

After presenting an introduction of the research work in previous Chapter, this chapter presents a comprehensive literature review of the state of the art literature. The chapter begins with the general literature review of anti-jamming communications in cognitive radio networks followed by game-theoretic literature review on anti-jamming communications in CRN. Finally, deception based anti-jamming techniques are discussed at the end to conclude the chapter. Moreover, two tables present the comparison of the literature discussed in this chapter.

2.1 Anti-Jamming in CRN

The CR is more vulnerable to security threats than other networks, such as cellular networks, due to its large scale and diversified environment. As a result, the CR's security requirements will be more restrictive than those of traditional wireless systems. Indeed, in the

absence of robust security solutions, attacks and CR malfunctions may outweigh any benefits. The security of a CR system is vulnerable to a wide range of attacks due to its large attack surface, including malignant radio jamming and denial of service (DoS). Different types of jammers discussed in the literature [61], namely, random jammer [62], constant jammer [18], reactive jammer [63, 64], sweep jammer [65, 66] and intelligent jammer [29] [67].

2.2 Anti-Jamming Games in CRN

In wireless communications, game theory is often utilised to tackle communication challenges such as resource allocation [55–57], packet relaying [36], and anti-jamming communication [29, 48, 58–60].

The dynamic interaction between legitimate users and the jammer has been extensively modeled and analysed using game-theoretic approaches. These anti-jamming games could be based on power domain [64, 68–74], code domain [18, 75], frequency domain [29, 30, 65, 66, 76–82], and space domain [83–86], as shown in Figure 2.1.

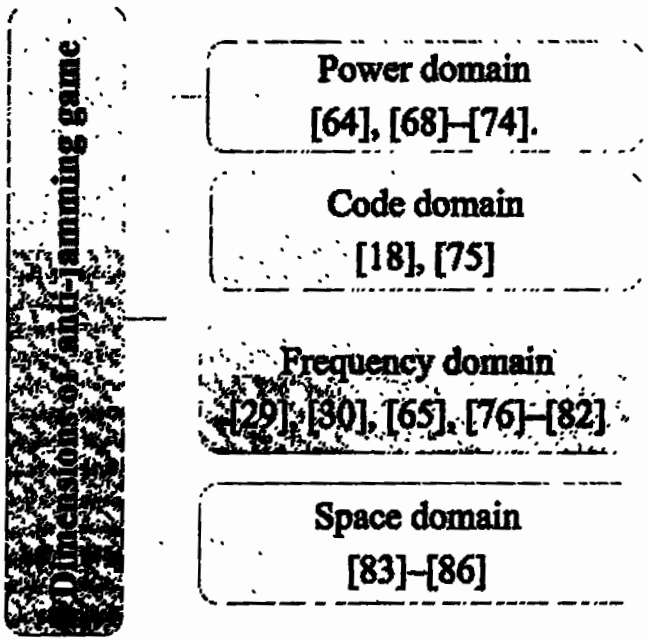


Figure 2.1: Dimensions of the anti-jamming game in cognitive radio networks.

2.2.1 Different Techniques of Anti-Jamming Games

Power Control Games

Anti-jamming power control (PC) communication has been done extensively using game-theoretic analysis [64, 68–74].

For instance, in [68], a power control Stackelberg game was presented as a leader follower game for jamming defence in cognitive radio networks. The problem is divided into sequential subproblems, follower sub-game, and leader-sub game. Another Stackelberg game was used in [69] for relay selection for the security of physical layers in cognitive radio networks. More specifically, the One Leader One Follower Stackelberg Game (OLOFS) was modelled to achieve optimal pricing strategy and power allocation in the pres-

Table 2.1: Summary of literature regarding types of anti-jamming games, jammer type, algorithm used, their defence techniques and equilibrium solutions.

Ref.	Game type	Jammer type	Solution method	Defense FH/RA/PC	NE/SE	Comments
[1]	Bimatrix game	Intelligent jammer	Q learning	FH	NE	FH solution is derived based on Q learning.
[2]	Zero-sum Markov game	Sweep jammer	Value iteration	FH+RA	NE	The constrained optimization problem is solved using value iteration
[68]	Stackelberg game	Intelligent jammer	Q learning	FH, PC	SE	The brief overview presented
[30]	Colonel Blotto game	Intelligent jammer	MLE based Q learning	FH, PC	NE	Adversarial Colonal Blotto game is analyzed
[64]	Stackelberg game	Smart jammer	Convex optimization	PC	SE	Stackelberg game basics
[87]	Colonel Blotto game	Intelligent jammer	Evolutionary game	PC	NE	Mixed strategy Nash equilibrium for Colonel Blotto game is derived.
[88]	Networked Colonel Blotto game	Malicious jammer	Co-evoluation algo	PC	NE	Mixed strategy NE is achieved.
[70]	Stackelberg game	Intelligent jammer	Scalable decomposition algo	PC	SE, NE	Existence of NE is proved.
[89]	Stackelberg game	Reactive jammer	Throughput analysis	FH	SE	Channel hopping time is analyzed in an adversarial environment
[80]	Bimatrix zero sum game	Pseudo random jammer	Quadratic programming	FH	NE	The constrained bimatrix FH game is solved using the quadratic programming.
[90]	Stackelberg game, Bayesian game	Intelligent jammer	Equilibrium analysis	PC	NE, SE	Equilibrium analysis of incomplete information games
[91]	Stackelberg game	Friendly jammer	Convex optimization	FH	SE	Artificial noise bsed jamming scheme is proposed
[92]	Stackelberg game	Reactive jammer	Equilibrium analysis	PC	NE,SE	PC game analysis
[93]	PT based game	Smart jammer	Prospect theory	PC	NE	Prospect theory-based game is presented and NE is analyzed
[94]	Non zero sum game	smart jammer	Water filling principle	PC	NE	Umqueness and existence of NE is proved.
[95]	Stackelberg game	Intelligent jammer	Multi-armed bandit (MAB)	PC,FH	SE	Multi-domain anti-jamming scheme is presented.
[67]	Bimatrix game	Intelligent jammer	Q learning	FH	NE	MDP based q learning is presented
[29]	Stochastic game	Intelligent jammer	Minmax Q	FH	NE	Zero sum game is solved using minmax q learning

ence of two eavesdroppers. Furthermore, the Primary User (PU) and the selected relay simultaneously achieve a Nash Equilibrium (NE).

In [70], the authors presented an adaptive approach to defend the jamming attacks in CRN by controlling transmission powers of the nodes, where the network topology is adaptively updated to nullify the effects of the jammer. The trade-off between jamming immunity and network coverage is seen as an optimization problem, which

can be solved by scalable decomposition strategies. The authors also present a continuous version of the game by considering continuous action space for both players. The authors in [71] study PC games for multi-user communication to combat jamming. Stackelberg's relay selection is used in [96] for physical layer security in cognitive radio networks. The Stackelberg Single-player Follower Game (OLOFS) is designed to achieve optimal pricing strategy and energy allocation in the presence of two eavesdroppers. The primary source and the selected relay operate simultaneously to achieve a Nash Equilibrium (NE). For the class of two person zero sum games, the Stackelberg equilibrium (SE) is also a NE.

Frequency Hopping Games

Game theory has also been used to investigate the frequency hopping anti-jamming communication in wireless communication networks. For instance, anti-jamming communication in CRNs with unknown channel statistics has been studied in [97]. The authors formulated the problem of anti-jamming multi-channel access in CRN as a non-stochastic multi-armed bandit problem, where both secondary sender and receiver choose their common operating channels by minimizing the probability of being jammed. Another interference avoidance-based channel-hopping stochastic game was investigated in a multi-agent environment in [65], where game-theoretic

based reinforcement learning mechanism is used to avoid jamming.

The authors in [1] recently presented a bandwidth-efficient frequency hopping game in wireless sensor networks. The authors in [98] presented a brief overview of anti-jamming communication in the context of dynamic spectrum access. Two typical ways of thwarting jammers are adaptation of transmission rate and Frequency Hopping (FH). These two are jointly adapted by [2] to improve the average throughput and provide better jamming resiliency against reactive sweep jammer. Specifically, the interaction between the jammer and the legitimate user is modelled in [2] as a Zero-Sum Markov Game (ZSMG), and a constrained NE is derived. The authors in [78] utilized a game theoretic framework to access an optimal channel in the presence of attacker, hence maximizing the channel payoff.

The channel Hopping (CH) based rendezvous scheme is adapted for the SU to meet and make the connection for further communications [77, 99]. This bounded time rendezvous scheme neither uses pre-shared secrets nor is role pre-assignment needed for bringing the SUs on a commonly available channel.

In [79], the authors presented a mobility-based Single Leader and Multiple Follower Stackelberg game (SLMFSG) to avoid jamming for increasing the network life in the WSN. Anti-jamming games in multi-channel cognitive radio networks were presented in [30], where the SU hops to another channel to avoid the jamming.

A zero-sum game is played between the attacker and the SU based on a Markov Decision Process (MDP). Maximum Likelihood Estimation (MLE) and Q learning are used for SU to learn from their environment.

In [76], the authors proposed a Hierarchical Learning Algorithm (HLA) for anti-jamming channel selection strategies in the presence of co-channel interference as a Stackelberg game. They considered jammer and users as independent learners (ILs), which choose their strategies independently and selfishly.

In [80], anti-jamming FH game is constructed using a bi-matrix game between the jammer and the legitimate user. In [81], game-theoretic stochastic learning approach is used for anti-jamming communication in dense wireless networks. In [66], the authors have considered joint multi-agent learners in stochastic game settings against a sweep jammer. They presented a collaborative multi-agent anti-jamming algorithm based on reinforcement learning in wireless networks. Markov game is formulated to model and analyze the anti-jamming problem in multi-user environment. Moreover, A Novel Distributed Multi-Agent Reinforcement Learning Algorithm against Jamming Attacks are presented in [100, 101].

Time domain countermeasure against random pulse jamming using MDP and reinforcement learning was presented in [62]. In [65], MARL is used as independent Q learning for each agent against a

sweep jammer as a common practice.

Another game-theoretic anti-jamming scheme for CRN is presented in [67], where the SU used Q learning to learn the dynamics of the jammer and reduce the complexity of value iteration-based learning. This scheme is considered as the benchmark scheme. However, they only consider the anti-jamming in ideal channel conditions with no noise present. Secondly, they did not consider the time variations in the wireless channel. The framework presented in [67] is improved by considering time-varying variable channels, a more realistic CRN approach. Furthermore, in present situation, both players' utility is reliant on channel quality; the better the channel, the bigger the reward, and vice versa. The sub-channels are differentiated based on the received SNR, which results in the varying maximum channel capacities.

Evolutionary game theory (EGT) has captured the attention of researchers in DSA because of its impressive ability to model heterogeneous environments as an evolving game. Evolutionary game theory is also lucrative because it relaxes the traditional rationality assumptions of game theory, which require all players to have complete knowledge of the game. Yet another advantage of EGT is that its framework of replicator dynamics can provide commutable rates of convergence to an Evolutionary Stable Strategy (ESS), thus generating concrete predictions of the distribution of the deployed

strategies and a picture of the adaptation of users over time [102].

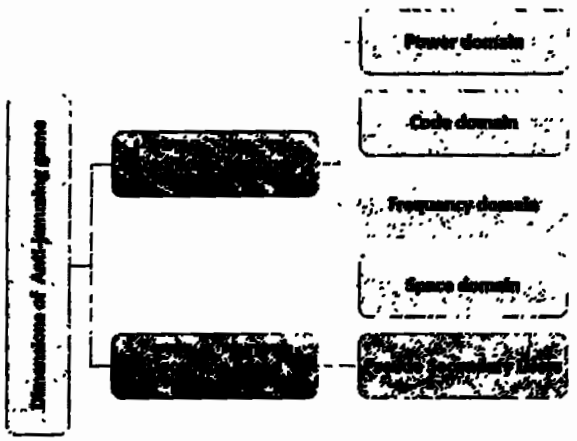


Figure 2.2: Dimensions of the anti-jamming game with an emphasis on deception-based techniques.

2.3 Deception-based Anti-Jamming Games

Deception empowers network administrators by thoroughly defending against attacks from both external parties and hostile insiders, properly warning when something is wrong, and offering precise threat intelligence for quick remediation. When compared to traditional anti-jamming schemes, deception give more protection by making the jammer waste its resources as shown in Figure 2.2.

Ahmed et. aln provides. in [60] used Stackelberg game based deception strategy against a deceiving jammer in CRN. The authors have used honeypots to detect deceiving jammers and used the jammers direction of arrival to place the jammer's direction in the nulls of the antenna. However, the proposed work is different from [60], because a deception strategy is used not only to detect the intruder but

also to deceive the jammer. The authors of [103] presented deceptive attack and defence game in honeypot-enabled networks for the Internet of Things (IoT). The authors analysed the deceptive attack and defence mechanism using game theory as a dynamic Bayesian game for single-shot and repeated games.

Recently, Nan et. al. [104] presented a leader-follower Stackelberg deception game based on power allocation. They considered two pairs of transmitter-receiver. The objective of the defender transmitter-receiver pair is to maximize the throughput of legitimate transmitter-receiver pair by deceiving the jammer with another transmitter-receiver pair. The jammer divide its limited power budget into two communication channels, hence reduces the power injected to the legitimate transmitter-receiver pair that transmit the real information. They also presented the sub-game perfect Nash Equilibrium (SPNE) of the deception game. The authors in [86] proposed a defensive defence against reactive jamming attacks in a communication channel. The tranceiver node adjusts its power levels hence modifying the real-time information intentionally, resulting in asymmetric uncertainty to decoy the adversary. Similarly, Hoang et. al. [105] presented a deception strategy against a reactive jammer using energy harvesting and backscatter technologies. The authors in [106] uses a deception strategy in which the transmitter uses an intelligent deception method in which it emits fake signals

Table 2.2: Comparison of the literature as evidence of novelty presented in this work.

Ref.	Deception	TS	FH	RA	PC	EH / BS	Throughput	Intelligent Jammer
[1]			✓				✓	✓
[2]			✓	✓			✓	
[67]			✓					✓
[29]			✓				✓	✓
[109]				✓			✓	✓
[60]	✓					✓		
[103]	✓							
[86]	✓	✓			✓			
[105]	✓					✓		
[104]	✓				✓			
[108]	✓	✓						✓
Proposed scheme	✓			✓			✓	✓

in order to attract the jammer. The transmitter then has the option of harvesting energy from the jamming signals or backscattering the jamming signals to broadcast data. As a result, jamming signals can help increase average throughput and decrease packet loss. Furthermore, defensive deception using game theory and machine learning is thoroughly reviewed and summarized in [107].

Moreover, the proposed work is different from [60] in the sense that a deception strategy is used not to detect the intruder but to deceive the rate-aware intelligent jammer. In contrast to what is suggested by the authors in [108], which proposed a queuing-based deception mechanism in CRN, we suggest a novel physical layer-based deception technique with the freedom to adapt the rate to the target parameter in CRN. The novelty of proposed approach is evident from the comparison as shown in Table 2.2.

CHAPTER 3

Game-Theoretic System and Adversary Models

After presenting literature review in previous chapter, this chapter presents a geme-theoretic system and adversary models to be used further in the next chapters for game-theoretic analysis. Three jammers with different levels of intelligence are discussed followed by system and jammer model for problem I and II, respectively.

The objective of the secondary user is to carefully switch the channels to maximize the spectrum utilization while avoiding the potential jamming. On the other hand, the jammer aims to forbid secondary users to form effective channel utilization by strategic jamming as shown in Figure 3.1. The objectives of the two players, jammer and secondary user, of the game, are opposite and there is no question of coordination. Therefore the dynamic interaction between them is well formulated as a non-cooperative zero-sum game,

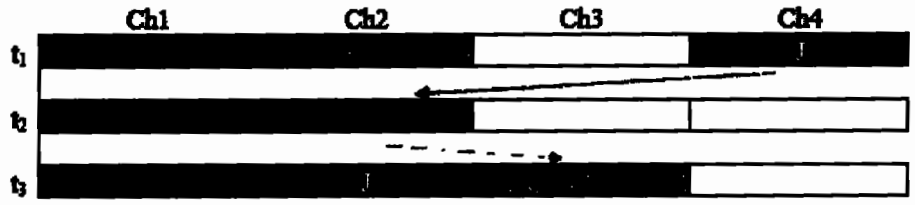


Figure 3.1: SU and jammer continuously hop their sub-channels to meet their objectives. SU would like to hop to other available sub-channels to avoid jammer.

where the gain of one player is the loss of other players.

It is noteworthy to mention that all secondary and pseudo secondary users coordinate to maximize their payoff therefore, all secondary users can be deemed as one player, and on the other hand, all attackers are considered to be as another player. Spectrum availability, channel quality and strategies of both secondary users and jammers are assumes time-varying. Players of the game are intelligent and hop heterogeneous in search of optimal space to avoid jamming. Cooperation cannot be taken for granted as the two participating players are opponents to each other and a gain of one player is the loss of other players.

3.1 Jammer Model With Different Levels of Intelligence

Jamming is a hostile attack in the CRN, where it disrupts the wireless communication by generating high-power noise at the targeted sub-channel as shown in Figure 3.2. A jammer with multiple intel-

ligence levels is considered.

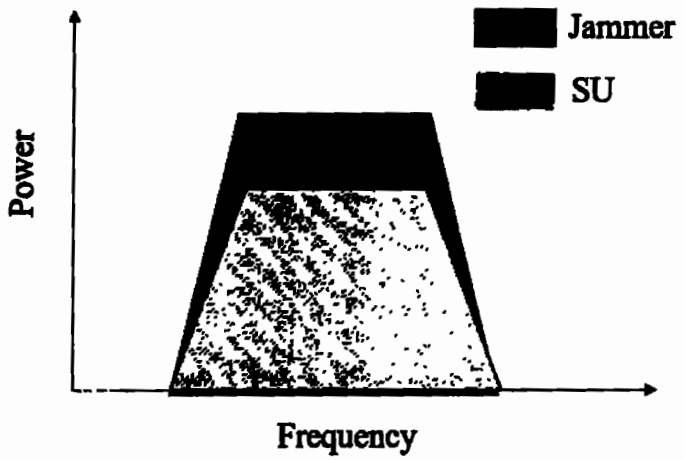


Figure 3.2: Deliberate radio jammer disrupts wireless communication by generating high-power noise at the targeted sub-channel.

3.1.1 Level-0 Intelligence: Random/ Infant Jammer

In level-0 intelligence, the jammer is considered as random in deciding which channel to jam.

3.1.2 Level-I Intelligence: Reactive / Baby Jammer

In level-I intelligence, it is anticipated that an intelligent jammer with cognitive skills will choose the optimal approach in response to channel dynamics and SU strategies. It scans its RF environment and transmit the jamming signals only if it finds the SU, hence saving its power.

3.1.3 Level-II Intelligence: Smart / Teen Jammer

In level-II intelligence, an intelligent jammer targets the highest impact communication (HIC) channels by targeting specific transmission characteristics of SU e.g., it can target the highest transmission rate $\mathcal{R}$, which may be the case in multimedia communications, highest transmission power $\mathcal{P}$, the channel with the highest bandwidth $\mathcal{B}$, packet inter-arrival time, and frequency shift, etc. [108]. The highest impact communication is quantitatively measured by jammer using attraction factors in Equation (3.3.1). The jammer perceives highest impact communication as the communication with the highest transmission rate $\mathcal{R}$, which is the case for multimedia communications.

3.2 System Model of Problem-I

The interweave paradigm for the time-slotted system is assumed in CRN, where SU can access the spectrum only if it is not used by the PU [12]. Every user scans the available sub-channels and starts transmission after white space is found.

We assume the network contains $\mathcal{H}$ PUs, $\mathcal{N}$ SUs, and $\mathcal{M}$ jammers. The channel's total used bandwidth is $\mathcal{W}$, and the whole bandwidth is split into $\mathcal{L}$ independent sub-channels of equal bandwidth $\mathcal{W}/\mathcal{L}$. However, the channel capacity $C_l(a^n, c_j^m)$ of each sub-channel

may differ, depending on the SNR of the received signal strength.

It is assumed that the jamming attack is the only source of channel deterioration in the network and any other source of interference, including the effects of multipath fading is neglected. Furthermore, perfect time and frequency synchronization between all SUs is assumed as in [110, 111]. Each sub-channel can be in two states, namely the idle state and the busy state. The relationship between the PU and an SU can be described by one of the two possible states of the sub-channel as follow:

- IDLE STATE: The channel is idle if any PU is not using it. The SU and the jammer are allowed to utilize an idle channel. The idle state of the sub-channel is represented by $\mathcal{P} = 1$
- BUSY STATE: The channel is considered busy if any PU is using it. Both SU and jammer are not allowed to transmit over a busy channel. This state is represented by $\mathcal{P} = 0$.

The channel states (idle or busy) are not known before the sensing action is taken place. In this work, $\mathcal{L}$ sub-channels are considered , where the quality of each sub-channel is different. Each sub-channel has its maximum capacity limit based on its received SNR, given by:

$$C_{l,t}(a^n, c_j^m) = \frac{\mathcal{W}}{\mathcal{L}} \log_2(1 + SNR_{l,t}^n(a^n, c_j^m)), \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M} \tag{3.2.1}$$

where $C_{l,t}(a^n, c_j^m)$ represents the capacity of the l^{th} sub-channel for the n^{th} SU at time slot t, where a^n and c_j^m are the actions of the

n^{th} SU and the m^{th} jammer, respectively. Moreover, $\mathcal{W}/\mathcal{L}$ is the bandwidth in Hz for each of the $\mathcal{L}$ sub-channels and $\text{SNR}^n_{l,t}(a^n, c^m_j)$ is the received SNR of the l^{th} channel for the n^{th} SU. Let us first consider the case where there is no jammer present in the system and the SNR is defined as:

$$\text{SNR}^{\text{n}}_{\text{l,t}}(\text{a}^{\text{n}}, \text{c}^{\text{m}}_{\text{j}}) = \frac{\mathcal{P}^{\text{n}}_{\text{l,t}}}{\mathcal{N}_{\text{o}}.\mathcal{W}/\mathcal{L}}, \quad \forall \text{n} \in \mathcal{N}, \forall \text{m} \in \mathcal{M} \tag{3.2.2}$$

where $\mathcal{P}^n_{l,t}$ is the average signal power received by the n^{th} SU at the l^{th} sub-channel at time slot t and $\mathcal{N}_o$ is the power spectral density (PSD) of the Additive White Gaussian Noise (AWGN). A high $\text{SNR}^n_{l,t}(a^n, c^m_j)$ would give a high channel capacity $\mathcal{C}_{l,t}(a^n, c^m_j)$ and hence a higher channel quality. The channel capacity of the n^{th} SU at the l^{th} sub-channel can be computed as:

$$\mathcal{C}_{l,t}(a^n, c^m_j) = \frac{\mathcal{W}}{\mathcal{L}} \log_2(1 + \frac{\mathcal{P}^n_{l,t}}{\mathcal{N}_o.\mathcal{W}/\mathcal{L}}), \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M}. \tag{3.2.3}$$

Moreover, the Signal to Interference plus Noise Ration (SINR) in the presence of a jammer can be calculated as $\text{SINR}^n_{l,t}(a^n, c^m_j) = \frac{\mathcal{P}^n_{l,t}}{\mathcal{N}_o.\mathcal{W}/\mathcal{L}+\mathcal{N}_{j,l}\mathcal{B}_j}, \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M}$ where $\mathcal{N}_{j,l}$ is PSD of the jamming signal and $\mathcal{B}_j$ is the bandwidth of the jammed channel. Since all sub-channels have identical bandwidth of $\mathcal{B}_j = \frac{\mathcal{W}}{\mathcal{L}}$, the SNR becomes:

$$SNR^n_{l,t}(a^n, c^m_j) = \frac{\mathcal{P}^n_{l,t}}{\mathcal{N}_o.\mathcal{W}/\mathcal{L} + \mathcal{N}_{j,l}\mathcal{W}/\mathcal{L}}, \forall n \in \mathcal{N}, \forall m \in \mathcal{M}. \tag{3.2.4}$$

The equation Equation (3.2.4) is true only when $a^n = c^m_j$, i.e., both the SU and the jammer are on the same channel. This results in severe degradation of the SNR for the SU. The objective of the SU is to

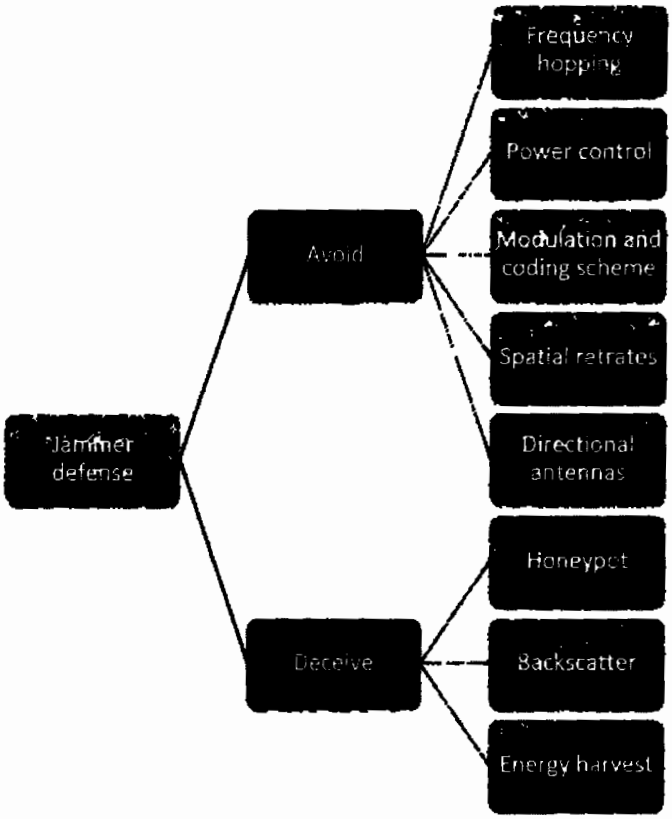


Figure 3.3: Taxonomy of defence techniques against a hostile jammer in CRN.

carefully switch to the available high-capacity channel to maximize the spectrum utilization, while simultaneously avoiding the potential jamming.

3.2.1 Secondary User Model-I

The SU senses its environment during its sensing period, before initiating any data transmission. Contention-based channel selection algorithm uses a structure called Preferable Channel List (PCL) to initiate data transmission. The algorithm avoids collision and performs Request To Send / Clear To Send (RTS-CTS) contention for

data transmission. Moreover, sensing-assisted access (SAA) protocol may be used as a complete random access mechanism for CRN to initiate data transmission. In this mechanism, the contention-based access is designed based on integrating the backoff process and spectrum sensing [112, 113]. However, in this contribution, this aspect is not investigated. During the sensing period, each SU would try to sense for the presence of any PU in the available sub-channels. However, the SU can not detect the presence of a jammer at the beginning of the time slot. Nonetheless, the SU can realize the presence of the jammer at the end of each time slot. More specifically, at the end of each time slot, the SU would know if its transmission was successful or was jammed by a malicious jammer. The interested readers may refer to [45, 46] for details concerning jamming detection. A successful transmission yield a positive payoff to the SU, while a jammed transmission would yield a negative payoff to the SU. The utility of the n^{th} SU in the l^{th} sub-channel based on the actions of SU (represented as a^n) and the action of the jammer (denoted as c_j^m) at time slot t can be derived as:

$$\mathcal{U}_{l,t}^n(a^n, c_j^m) = \mathcal{C}_{l,t}^n(a^n, c_j^m)(x_{l,t}(a^n, c_j^m)(\mathcal{T}_l^n - \mathcal{E}_l^n) - (1 - x_{l,t}(a^n, c_j^m))(\mathcal{J}_l^n + \mathcal{E}_l^n)), \quad (3.2.5)$$

where $\mathcal{E}_l^n$ is the cost of transmission of the n^{th} SU, $\mathcal{T}_l^n$ is the SU gain factor for successful communication, $\mathcal{J}_l^n$ is the loss factor for

SU when the SU transmission in the l^{th} sub-channel is jammed. The impact of the sub-channel SNR and the decision of each player on the utilities of both players are shown in Figure 3.4. The received SNR at each sub-channel increases from sub-channel 1 to sub-channel 10. The utilities earned by each player are opposite to each other. The missing utility at sub-channel 4 in Figure 3.4 indicates that the PU is transmitting in this sub-channel, and neither SU nor jammer can use this sub-channel. Furthermore, sub-channel 9 was jammed in the previous time slot, if the jammer stays there, the SU would have negative utility to use sub-channel 9. Similarly, for the jammer, if SU was at sub-channel 9 and if it stays there, then the jammer would have positive utility at sub-channel 9. Additionally, the PU could change to a different sub-channel in each time slot, but it is assumed that both the SU and the jammer can detect the sub-channel used by the PU.

The utility of the proposed system in Equation (3.2.5) can be compared with the utility function in of the benchmark system [67] given by $G(s,a) = \sum_{l=1}^{L} G_l(s,a)$, where the gain of the SU at the l^{th} sub-channel is computed as:

$$G_l(s,a) = x_l(s,a) \times U - y_l(s,a) \times C, \tag{3.2.6}$$

while $x_l(s,a)$ and $y_l(s,a)$ are binary switching functions. Furthermore, U and C denote the utility earned by the SU and the jamming cost of SU, respectively. The authors in [67] assume that the values

for the utility and the cost in every sub-channel are identical, since their sub-channels have the same quality. The left-hand side of both Equation (3.2.5) and Equation (3.2.6) denotes the utility function of the SU, although different notations were used. The right-hand side of Equation (3.2.5) and Equation (3.2.6) has the following differences:

- The factor $C_{l,t}^{n}(a^n, c_j^m)$ is introduced in Equation (3.2.5) to differentiate the sub-channels based on the sub-channel capacities. Hence, successful transmission in a good quality sub-channel would yield better utility for the SU. This quality factor is missing in [67] and Equation (3.2.6).

- The factors $\mathcal{E}_l^n$ and $\mathcal{E}_{jl}^m$ are also considered to account for the transmission costs for SU and for the jammer, respectively, in terms of the battery utilization and the power transmitted.

- The two binary switching functions $x_l(s, a)$ and $y_l(s, a)$ in Equation (3.2.6), are used in [67] such that $x_l(s, a) + y_l(s, a) = 1$. To simplify the mathematical notation, the only one binary switching function $x_{l,t}(a^n, c_j^m)$ is used instead of two, such that $x_{l,t}(a^n, c_j^m) + (1 - x_{l,t}(a^n, c_j^m)) = 1$.

- Furthermore, a more detailed utility function for the jammer is considered compared to that of [67], as will be explained later in Equation (3.2.11) and Equation (4.1.2). Combining Equation

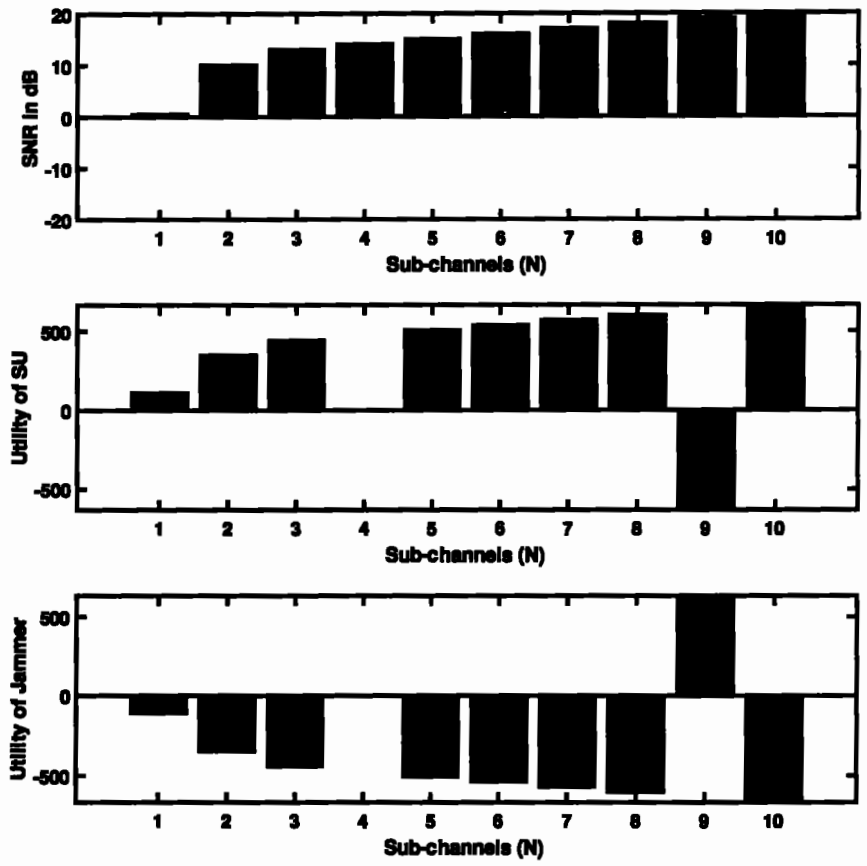


Figure 3.4: The impact of change in SNR and the decision of each player on the utilities of other players.

(3.2.5) and Equation (3.2.3) gives Equation (3.2.7).

$$\mathcal{U}_{l,t}^{n}(a^n, c_j^m) = \frac{\mathcal{W}}{\mathcal{L}} \log_2(1 + \frac{\mathcal{P}_{l,t}^{n}}{\mathcal{N}_o.\mathcal{W}/\mathcal{L}})(x_{l,t}(a^n, c_j^m)(\mathcal{T}_l^n - \mathcal{E}_l^n) - (1 - x_{l,t}(a^n, c_j^m))(\mathcal{J}_l^n + \mathcal{E}_l^n)), \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M}, \quad (3.2.7)$$

where $x_{l,t}(a^n, c_j^m) \in \{1, 0\}$ is a binary switching function used to

indicate successful/jammed SU communication:

$$x_{l,t}(a^n, c_j^m) = \begin{cases} 1, & a^n \neq c_j^m, \forall n \in \mathcal{N}, \forall m \in \mathcal{M}, \\ 0, & a^n = c_j^m, \forall n \in \mathcal{N}, \forall m \in \mathcal{M}. \end{cases} \tag{3.2.8}$$

Note that $x_{l,t}(a^n, c_j^m)$ is 1 for successful SU transmission and is 0 for jammed SU transmission. Specifically, $C_{l,t}^n(a^n, c_j^m) = 0$ if the SNR is below a certain threshold value SNR_{th} i.e., $\text{SNR}_{l,t}^n(a^n, c_j^m) \leq \text{SNR}_{th}$ and the value of the switching function $x_{l,t}(a^n, c_j^m)$ would also become 0. Equations (3.2.2), (3.2.4) and (3.2.8) are related in the sense that:

$$x_{l,t}(a^n, c_j^m) = \begin{cases} 1, & SNR_{l,t}^n(a^n, c_j^m) > SNR_{th}, \\ 0, & SNR_{l,t}^n(a^n, c_j^m) \leq SNR_{th}, \end{cases} \tag{3.2.9}$$

and

$$SNR_{l,t}^n(a^n, c_j^m) = \begin{cases} \frac{\mathcal{P}_{l,t}^n}{\mathcal{N}_o \mathcal{W}/\mathcal{L}}, & x_{l,t}(a^n, c_j^m) = 1 \\ \frac{\mathcal{P}_{l,t}^n}{\mathcal{N}_o \mathcal{W}/\mathcal{L} + \mathcal{N}_{J,l} \frac{\mathcal{W}}{\mathcal{L}}}, & x_{l,t}(a^n, c_j^m) = 0. \end{cases} \tag{3.2.10}$$

In other words, the SU utility function in Equation (3.3.3) incorporates the practical channel condition in terms of both the channel capacity and the jamming conditions. The goal of the SU is to maximize the expected sum of the discounted payoff by choosing a good quality channel that is not jammed by the jammer.

3.2.2 Jammer Model-I

Jamming is a hostile attack in the CRN, where it disrupts the wireless communication by generating high-power noise at the targeted sub-channel as shown in Figure 3.2. This research considers two

types of jammers, i.e., random jammer (level-0 jammer) and intelligent jammer (level-1 jammer). A random jammer would randomly jam a sub-channel in a different time slot. Inspired by [67] and [29], when an intelligent jammer with cognitive capabilities is assumed, it adapts the best strategy to observe the channel dynamics and the SU strategies. The jammer senses the RF environment for a given sensing duration and then transmits its jamming signals based on the channel conditions and the strategy of the SU. If a PU is detected in a sub-channel, the jammer would switch to other available sub-channels to avoid the heavy penalty imposed by law-enforcement agencies and start sensing again [29]. The utility function of the jammer in the l^{th} sub-channel is based on the actions of the SU and of the jammer, which is represented by:

$$\mathcal{U}^m_{jl,t}(a^n, c^m_j) = C^m_{l,t}(a^n, c^m_j)[(1 - x_{l,t}(a^n, c^m_j))(\mathcal{T}^m_{jl} - \mathcal{E}^m_{jl}) - x_{l,t}(a^n, c^m_j)(\beta \mathcal{E}^m_{jl})], \forall n \in \mathcal{N}, \forall m \in \mathcal{M}, \quad (3.2.11)$$

where $C^m_{l,t}(a^n, c^m_j)$ is the channel capacity of the l^{th} sub-channel and $\mathcal{T}^m_{jl}$ is the jammer gain factor when a SU was successfully jammed, while $\mathcal{E}^m_{jl}$ is the cost of transmitting the jamming signals. Furthermore, β is the jammer regret factor when the jamming was not successful, which is the negative reward earned by the jammer when the jammer sends a jamming signal to a sub-channel that the SU did not use. As mentioned in Equation (3.2.8), $x_{l,t}(a^n, c^m_j)$ is a switching

function having $x_{l,t}(a^n, c_j^m) = 0$, when the jammer successfully jams a channel (zero regrets), while $x_{l,t}(a^n, c_j^m) = 1$ when the jammer fails to jam the SU. Hence, an intelligent jammer is more inclined to jam a SU that operates in a high-capacity sub-channel than a low-capacity sub-channel. The objective of the jammer is to maximize the probability of successful jamming.

3.3 System Model of Problem-II

The system model for the first portion of the work described in this thesis was presented in the previous section. The second portion of the system model for the deception-based anti-jamming mechanism is shown here. Let's look at the jammer features as defined by the jammer model before discussing the system model to deceive the intelligent jammer.

3.3.1 Jammer Model-II

The objective of the Jammer: Jammer tries to minimize the average bandwidth efficiency of the CBS by injecting noise to those users having the highest impact communication. The impact of jamming attack on the communication sub-channel is to reduce the SNR at the receiver and hence reducing the capacity of the channel. A powerful jammer with the following characteristics is assumed.

- Level I: Reactive jammer
- Level II: Intelligent jammer: Rate-aware

The basic assumptions about the jammer model are:

- Jamming at the physical layer is assumed, in which a jammer hinders wireless communication by producing high-power noise at the targeted sub-channel.
- Similar to other secondary users, an intelligent jammer is in essence a secondary user with sensing, perception, and adaptive capabilities as assumed by [67, 114] and [30]. As we already know that the secondary user has the lower priorities to access the spectrum as compared to the primary users, therefore a secondary user has to use its sensing ability to sense its environment for the availability of free spectrum to continue its transmission and to avoid interference to the primary user. In a very similar fashion, a jammer needs to sense its RF environment for the presence of a primary user. Then the jammer will avoid the primary user if detected as the primary user due to the risk of a high penalty. On the other hand, the jammer targets users other than primary users.

  We are combating against intelligent jammer in CRN in interview paradigm, where the spectrum hole is accessed by secondary user on opportunistic spectrum access (OSA) bases. The

intelligent jammer has the cognitive capabilities, means that it is capable of sensing the presence of primary user in the network. The intelligent jammer is deemed as a secondary users with negative intentions to disrupt the communication of other secondary users. In the initial sensing time of a time slotted system, intelligent jammer do also listen to the presence/ absence of primary user just as the other secondary users do. So, the intelligent jammer is capable of sensing the RF environment and based on the sensing results it transmit the noise signals on the frequencies vacated by a primary user and utilized by secondary user to jam the communication of secondary users. The jammer can not jam the primary user due to heavy penalty imposed by the law-enforcement agencies.

- Due to the prospect of a severe penalty, the jammer does not target the PU communication [29, 108]. The jammer monitors the RF environment for a predetermined amount of time before transmitting its jamming signals in accordance with sub-channel circumstances and the SU's strategy. If a PU is identified in one of the sub-channels, the jammer will switch to another sub-channel and begin detecting another SU.

- Jammer is constrained by the J_{max} number of users it can target. For the cause of simplicity $J_{max} = 1$.

- The jammer aim at jamming the users with Highest Impact Communication (HIC) [108]. In multimedia communication, the HIC is the one having higher data rates. Therefore, high data rates for SU increases the risk of getting jammed by intelligent jammer.

- Jammer use attraction factor δ, $0 \leq \delta_i \leq 1$, $\forall i \in \mathcal{N}$ to target (HIC), which is defined as

$$\delta_i = \frac{\mathcal{R}_i}{\sum_{i=1}^{N} \mathcal{R}_i}, \quad \forall i \in \mathcal{N} \quad s.t. \sum_{i=1}^{N} \delta_i = 1 \tag{3.3.1}$$

  where $\mathcal{R}_i$ is the data rate adapted by the i^{th} SU/PSUs. The intelligent jammer may acquire the rate/code/modulation information of SU/PSU using one of the following three ways: explicit rate information, modulation guessing and code guessing [52].

- Jammer calculates $\delta, 0 \leq \delta \leq 1$ for every detected signal in his environment according to Equation (3.3.1) and target the SU with the highest attraction factor.

Using level-I intelligence, the jammer will scan the environment and know the channel qualities and the transmitting SUs. Using level-II intelligence, the jammer will determine the highest impact communication of the secondary network using attraction factor in Equation (3.3.1).

A triangular anti-jamming deception game is presented be-

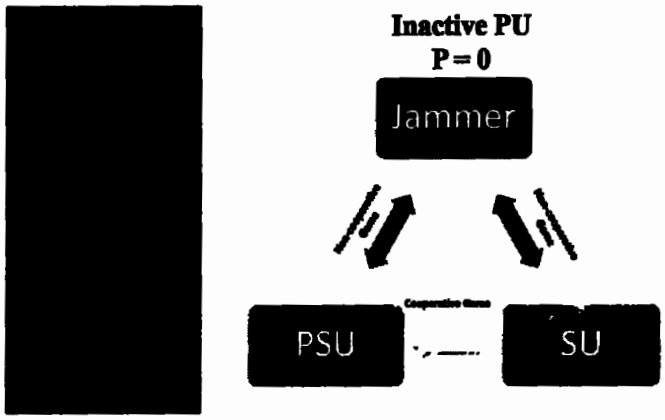


Figure 3.5: A triangular anti-jamming deception game is presented between SU, PSU, and the jammer in the absence of PU.

tween SU, PSU and the jammer in the absence of PU as shown in Figure 3.5, indicating a cooperation in the form of light green arrow between all SUs and the PSUs.

3.3.2 Cognitive Base Station (CBS)

The secondary network consists of one central unit called fusion center (FC) / Cognitive Base Station (CBS), several SUs and a PSU.

Objective of CBS: The goal of CBS is to improve the overall system's average throughput by successfully deploying a deception mechanism using PSU in the presence of an intelligent jammer.

3.3.3 Secondary User Model-II

- SUs are enforced by dynamic spectrum access (DSA) implemented in CBS to periodically pause its transmission and sense for PU activity, to protect PU incumbent services.

- The SU senses its environment during its sensing time before commencing any data transfer. Throughout the sensing time, each SU would try to detect the existence of any PU in the accessible sub-channels. On the other hand, the SU would be unable to identify the existence of an adversary at the start of a time slot. Concerning jamming detection, interested readers may see [46]. Despite this, the SU is able to detect the existence of the jammer at the ending of each time slot [45]. After each time slot, an SU would know if its communication was successful or was jammed by a malignant jammer.

- The SU would receive a positive reward if the transmission was successful, while a jammed transmission, on the other hand, would result in a negative payoff. The utility of the SU denoted as U_i in the i^{th} sub-channel can be derived as:

$$\mathcal{U}_i = \mathcal{C}_i((1-\delta_i)\mathcal{G} - \delta_i\mathcal{J}), \tag{3.3.2}$$

  where $\mathcal{C}_i$ and δ are defined in Equation (3.2.3) and Equation (3.3.1), respectively. Moreover, G and $\mathcal{J}$ are the gain of the secondary user and gain of the jammer, respectively. Combining Equation (3.3.2) with Equation (3.2.3) yields Equation (3.3.3).

$$\mathcal{U}_i = (\mathcal{W}/\mathcal{L})\log_2(1+\frac{\mathcal{P}_r}{\mathcal{N}_w.\mathcal{W}/\mathcal{L}})\{(1-\delta_i)\mathcal{G} - \delta_i\mathcal{J}\}. \tag{3.3.3}$$

In another way, the SU utility function in Equation (3.3.3) allows us to incorporate the practical channel condition both in terms of channel capacity and jamming situations. In the presence of a jammer,

the SU's objective is to achieve the expected sum of the discounted payoff by selecting a good quality channel.

3.3.4 Pseudo Secondary User (PSU) Model

A dummy user called pseudo secondary user (PSU) is assumed to mimic the characteristics of a legitimate SU to lure the jammer. PSU lures the jammer by transmitting at a higher modulation scheme to invite the jammer to attack the PSU. A group of SUs takes the services of a PSU to decoy the intelligent jammer.

- A PSU does not send a legitimate signal; rather, it transmits garbage data at a rate greater than the other SU to deceive the jammer. That is why the utility of PSU is not counted in the calculations of throughput of the CBS.
- The wireless communication system is designed to achieve a specific BER line in a BER-SNR(dB) curve. One has to follow this line for reliable communication. Moreover, according to the Shannon capacity theorem, a specific data rate can tolerate a certain level of jamming power (received SNR) to achieve reliable communication. If the jamming power is greater than the threshold, the corresponding data rate of transmission may exceed the channel capacity. The packets are lost, which results in lower throughput of the system. However, this does not apply

to the proposed PSU-based deception mechanism, as the PSU is transmitting garbage values with a rate higher than the rate of SUs in the vicinity. Since a PSU is transmitting garbage data, it is not interested in the loss of packets. Hence, the data rate/capacity of PSU is not calculated towards the throughput of the system.

- A variable called attraction factor $\delta \in (0, 1)$ is introduced to lure the jammer by attracting the rate-aware jammer towards itself. The proposed scheme is implemented so that the jammer gets the false impression of PSU as the highest impact communication. It becomes Achilles heel for the jammer.

3.4 Summary

The novel contributions of the research work can be summarized as follows:

1. A cognitive adversarial jammer is considered, which is an intelligent attacker that adapts to the dynamics of the sub-channels and the strategies of SU.
2. A mathematical modelling of the system is developed to satisfy the requirement, incorporating intelligence in the SU to cope with an intelligent jammer.
3. A more realistic and practical channel model is framed here,

which assumes that all sub-channels may have the varying quality of service. More specifically, the channel conditions may change over time and differ from one sub-channel to another.

4. The proposed framework considers various factors and parameters that capture the near practical channel dynamics, i.e., SNR, variable channel capacity, jamming gain, transmission cost and jamming cost of each player in the game.

CHAPTER 4

Part-I: Proposed Anti-Jamming Game to Combat Intelligent Jamming for Cognitive Radio Networks

The system model was presented in the previous chapter. This chapter describes the optimal solution strategy to combat intelligent jammers.

4.1 Game Theoretic Anti-Jamming Mechanism-I

Here the game-theoretic anti-jamming mechanism is presented. An anti-jamming game is presented after having a brief section of preliminaries of the game theory.

4.1.1 Preliminaries

A Stochastic Game (SG) is the natural extension and generalization of the Markov Decision Process (MDP) to multi-agent systems [9, 26].[1] SG provides a framework for multi-agents in multiagent reinforcement learning (MARL). In this contribution, a stochastic anti-jamming game is developed between two players of conflicting interests.

Definition 1 [27]: A two-player stochastic game is defined as $\mathcal{G} = \langle \mathcal{X}, \mathcal{S}, \mathcal{A}_i, \mathcal{U}_i \rangle$, where $\mathcal{X} = \{1, 2\}$ is the index of the players, $\mathcal{S}$ is the discrete strategy space of the game, $\mathcal{A}_i$ is the discrete action space of player i, while $\mathcal{U}_i : S \times A_i$ is the utility/payoff of player i, $\forall i \in \mathcal{X}$.

Definition 2[115]: A pair of strategies (i^*, j^*) $\{i$ for the row player, and j for the column player$\}$ yields a non-cooperative Nash equilibrium solution to a bimatrix game $(\mathcal{A} = \{A_{ij}\}\ \mathcal{B} = \{B_{ij}\})$, where $\mathcal{A}$ and $\mathcal{B}$ are payoff matrices for each player if the following two inequalities are satisfied:

$$\begin{aligned} A_{i^*j^*} &\geq A_{ij^*}, \quad \forall i, j = 1, 2, 3, ..., P, \\ B_{i^*j^*} &\geq B_{i^*j}, \quad \forall i, j = 1, 2, 3, ..., P, \end{aligned} \tag{4.1.1}$$

where P is the total number of pure strategies, each stage of a stochastic game can be viewed as a bimatrix game.

The basic assumption of a stochastic game between two inter-

[1] Stochastic games are also called Markov games, which essentially are n-agent MDP [27].

acting players is that the actions of each player will have an impact on the utility of other player. The same assumption is valid here in this research work. The SU obtained its utility based on the past actions of the jammer in the previous time slot. If the sub-channel to be accessed by the SU receives an SNR lower than a threshold, it implies that the jammer successfully jammed the sub-channel, and SU will get lower utility at that sub-channel at time slot t.

4.1.2 Game Formulation

Based on the knowledge about the channel, the system and the attacker, the objective of the SU is to carefully choose a sub-channel to maximize its spectrum utilization while avoiding jamming. On the other hand, the jammer aims to forbid the SU from effective channel utilization by a strategic jamming approach. The objectives of the two players, namely the jammer and the SU, are opposite to each other. Therefore, the dynamic interaction between them is well formulated as a non-cooperative game, where the gain of one player is the loss of another player. Furthermore, spectrum availability, quality of the channel, and strategies of both SU and jammer can be time-varying. Players are assumed to be intelligent and would exhibit rational behaviour to maximize their payoffs according to their objectives.

A two-player SG is formulated between the SU and the jam-

mer as described below:

Players: There are two non-cooperative players in the game namely the SU and the jammer.

States: Every sub-channel occupation is considered as the state S of the game. For example, if there are $\mathcal{L}$ sub-channels, then there are $\mathcal{L}$ states. The number of available states to the SU and jammer is given by $\mathcal{L} - \mathcal{H}$, where $\mathcal{L}$ is the total number of sub-channels and $\mathcal{H}$ is the number of PUs in the network.

Actions: An action $A(s)$ at each state has $\mathcal{L} - \mathcal{H}$ hopping possibilities. For each $\mathcal{L} - \mathcal{H}$ available states, the possible action set is $\mathcal{A}(s) = \{a_1, a_2, a_3, ..., a_i, ..., a_{\mathcal{L}-\mathcal{H}}\}$, where a_i is the action to hop to the i^{th} sub-channel in the $\mathcal{L}-\mathcal{H}$ available sub-channels. Both players choose actions to hop to any of the available sub-channels, which are not occupied by the PU. Since the available frequency slots are the same for both players, the action set $\mathcal{A}(s)$ is the same. Every action results in a change of state. Both jammer and SU sense the channel during the sensing period and hence the channel states and channel quality are assumed to be common knowledge in the game.

Payoff: The immediate payoffs of both players in a bimatrix game at each stage are given by Equation (3.3.3) and Equation (3.2.11), respectively. The total utility of the whole secondary network is given by

$$\mathcal{U}_{T,t}(\boldsymbol{a}, \boldsymbol{c}_j) = \sum_{n=1}^{\mathcal{N}} \sum_{l=1}^{\mathcal{L}} \mathcal{U}_{l,t}^{n}(a^n, c_j^m), \tag{4.1.2}$$

where $\mathcal{U}_{l,t}^{n}(a^n, c_j^m)$ is given by Equation (3.3.3). The long term objective of SU is to maximize the expected sum of discounted payoff, which can be written as [26, 29]:

$$\max_{\mathbf{a}} E\left\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{T,t}(a, c_j)\right\}, \tag{4.1.3}$$

where γ^t is a time decaying discount factor, $0 < \gamma^t < 1$, that determines the significance of future payoffs and $\mathcal{U}_{T,t}(a, c_j)$ is the utility of the secondary network at time t, which is given by Equation (4.1.2). The frequency hopping strategy of the SU is to maximize its utility by taking an optimal action that is given by:

$$a^* = \arg\max_{a\in\mathcal{A}(s)} E\left\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{T,t}(a, c_j)\right\}. \tag{4.1.4}$$

Similarly, the frequency hopping strategy of an intelligent jammer is to maximize its expected utility of $\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{jl,T,t}(a, c_j)$ by taking an optimal action of:

$$c_j^* = \arg\max_{c_j\in\mathcal{A}(s)} E\left\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{jl,T,t}(a, c_j)\right\}. \tag{4.1.5}$$

The pair (a^*, c_j^*) is said to be an equilibrium pair, if Equation (4.1.4) and Equation (4.1.5) follow the following inequalities:

$$\begin{aligned} E\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{T,t}(a^*, c_j^*)\} &\geq E\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{T,t}(a, c_j^*)\}, \\ E\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{jl,T,t}(a^*, c_j^*)\} &\geq E\{\sum_{t=0}^{\infty}\gamma^t \mathcal{U}_{jl,T,t}(a^*, c_j)\}. \end{aligned} \tag{4.1.6}$$

4.1.3 Q-learning-based Anti-Jamming

In the previous section, the anti-jamming game formulation was mentioned. This section describes the defense strategy of the SU, i.e.,

Algorithm 1 Game theoretic frequency hopping algorithm.

1: Initialize $k = 0$, $K = 100$, $\alpha^k \in [0,1]$ and $\gamma \in [0,1]$, $\forall s^k \in \mathcal{S}$, $a^k \in \mathcal{A}(s)$ and $c_j^k \in \mathcal{A}(s)$.
2: Let $Q(s^k, a^k) = 0$.
3: **while** $k \to K$ **do**
4: Execute action a^k, get immediate reward $R(s, a^k)$ according to (3.3.3) and observe s^{k+1}.
5: Choose a^{k+1} from s^{k+1} using policy derived from $Q(s^k, a^k)$ as given in (4.1.9).
6: Update $Q(s^k, a^k)$ for SU according to (4.1.7).
7: $s^k \leftarrow s^{k+1}$ and $a^k \leftarrow a^{k+1}$.
8: **end while**
9: **if** $a = c_j$ **then**
10: The channel is jammed.
11: **else**
12: The SU transmission is successful.
13: **end if**

how to defend the SU from being jammed by the jammer. A multi-agent reinforcement learning (MARL) agent as independent learner (IL) in Q-learning is used to combat jamming attacks in CRN. See Section 1.5 for details of reinforcement learning algorithm called Q-learning to be used here for solving the optimal strategy. The value of $Q(s^k, a^k)$ in Equation (1.5.2) can be updated recursively without having to estimate the transition probabilities [116], as follows:

$$Q(s^k, a^k) = Q(s^k, a^k)(1 - \alpha^k) + \alpha^k[R(s^k, a^k) + \gamma \max_{a^{k+1}}\{Q(s^{k+1}, a^{k+1})\}], \qquad (4.1.7)$$

where α^k is the linear learning rate satisfying $0 < \alpha^k < 1$. For α^k to be time decaying it must satisfy the conditions $\sum_{k=0}^{\infty} \alpha^k = \infty$ and $\sum_{k=0}^{\infty} (\alpha^k)^2 < \infty$. Note that γ is the discount factor satisfying $0 < \gamma < 1$. Similarly, the intelligent jammer can also adapt the Q-learning algorithm to learn the dynamics of the SU and the sub-channels [117]. Explicitly, the jammer maintains a separate Q-learning table for its own rational decisions. After updating the Q

Algorithm 2 Game theoretic algorithm for jammer.

1: Initialize $k = 0$, $K = 100$, $\alpha^k \in [0,1]$ and $\gamma \in [0,1]$, $\forall s^k \in \mathcal{S}$, $a^k \in \mathcal{A}(s)$ and $c_j^k \in \mathcal{A}(s)$
2: Let $Q_j(s^k, c_j^k) = 0$.
3: **while** $k \to K$ **do**
4: Execute action c_j^k, get immediate reward $R_j(s, c_j^k)$ according to (3.2.11) and observe s^{k+1}.
5: Choose c_j^{k+1} from s^{k+1} using policy derived from $Q_j(s^k, c_j^k)$ as given in (4.1.10).
6: Update $Q_j(s^k, c_j^k)$ for jammer according to (4.1.8).
7: $s^k \leftarrow s^{k+1}$ and $c_j^k \leftarrow c_j^{k+1}$.
8: **end while**
9: **if** $c_j = a$ **then**
10: The jammer is successful.
11: **else**
12: The SU transmission is successful.
13: **end if**

table, an intelligent jammer could decide based on its updated Q table. The update rule for the jammer's Q value is given by:

$$Q_j(s^k, c_j^k) = Q_j(s^k, c_j^k)(1 - \alpha^k) + \alpha^k[R_j(s^k, c_j^k) + \gamma \max_{c_j^{k+1}}\{Q_j(s^{k+1}, c_j^{k+1})\}], \quad (4.1.8)$$

where $Q(s^k, a^k)$ and $Q_j(s^k, c_j^k)$ are the estimates of the expected sums of the discounted payoffs for both the SU and the jammer, respectively, which could evolve. The rewards of the SU and the jammer after choosing their respective actions at state s^k are given by $R(s^k, a^k)$ and $R_j(s^k, c_j^k)$, respectively. These immediate rewards are calculated using Equation (3.3.3) and Equation (3.2.11) for the SU and the jammer, respectively. The SU would stay on the current sub-channel if its reward on the current sub-channel is good enough to contribute to Q value update positively. A negative instant reward in a certain sub-channel indicates that the sub-channel has been jammed and the SU should avoid that sub-channel by hopping

to another available sub-channel in the next round. The SU's optimal frequency hopping strategy to avoid the jammer is the action that maximizes its Q-value in state s and is given by:

$$a^* = \arg \max_{a \in \mathcal{A}(s)} Q(s, a). \tag{4.1.9}$$

Algorithm 1 summarizes the game-theoretic frequency hopping algorithm for SU. The jammer's optimal frequency hopping strategy to jam the SU is a greedy policy that chooses the action with maximum Q-value in state s and is given by

$$c_j^* = \arg \max_{c_j \in \mathcal{A}(s)} Q_j(s, c_j). \tag{4.1.10}$$

For jammer, the procedure is summarized in Algorithm 2.

4.1.4 Complexity Analysis

Inspired by [67], the computational complexity of Algorithm 1 and Algorithm 2, in this subsection is derived. Inside while loop line 5 of both Algorithms represent the policy derived from Q learning, and line 6 represent the update equation of Q learning. The computational complexity in each iteration of the policy phase comes from solving linear equations. The complexity of the policy phase is given by $O(|S|)$. On the other hand, the complexity of Q learning phase is calculated as $O(|A|.|S|)$. Combing both will result $O(|S|.(1+|A|))$. As the algorithms run for K number of iterations, hence the overall complexity may be represented by $O(|K|.|S|.(1+|A|))$, where S and A represent states and actions, respectively.

Table 4.1: Parameters used in the simulations, and different SNR values for all sub-channels in case I.

l	SNR_{dB}	Other Parameters	
1	0	$\mathcal{L}$	10
2	10	$\mathcal{H}$	1
3	13	$\mathcal{N}$	1
4	14	$\mathcal{M}$	2
5	15	$\mathcal{T}_{jl}$	-100
6	16	$\mathcal{E}_l$	10
7	17	$\mathcal{E}_{jl}$	10
8	18	$\mathcal{T}_l$	100
9	19	α	0.1
10	20	β	-100
	$\overline{\text{SNR}_{\text{dB}}} = 16.2$	γ	0.8

After presenting the system and adversary model and then the solution mechanism in the previous section, this section presents the results and discussions. Problem-I presented the results against random and reactive jammers while Problem-II provided deception-based anti-jamming results against rate-aware intelligent jammers.

4.2 Results and Discussions

This simulation study considers $N = 10$ sub-channels, one SU, one PU, and up to four jammers. When the PU occupies a sub-channel, neither the SU nor the jammers can access that channel. The SU chooses a high-capacity sub-channel that is potentially jamming-free, while the jammers predict and choose the sub-channel used by the SU.

4.2.1 The Effect of Using Different Channel Types:

The capacity of a sub-channel is a measure of the highest information rate that can be achieved with a very small error rate.The channel capacity $C_{l,t}^{n}(a^{n}, c_{j}^{m})$ is represented by Equation (3.2.1) and Equation (3.2.10), while the bandwidth efficiency in bits per second per Hertz (bps/Hz) can be computed as:

$$\eta^{n} = \frac{C_{l,t}^{n}(a^{n}, c_{j}^{m})}{\frac{\mathcal{W}}{\mathcal{L}}} = \log_{2}\left(1 + \frac{\mathcal{P}_{l,t}^{n}}{\mathcal{N}_{o}.\mathcal{W}/\mathcal{L}}\right), \quad \forall n \in \mathcal{N} \tag{4.2.1}$$

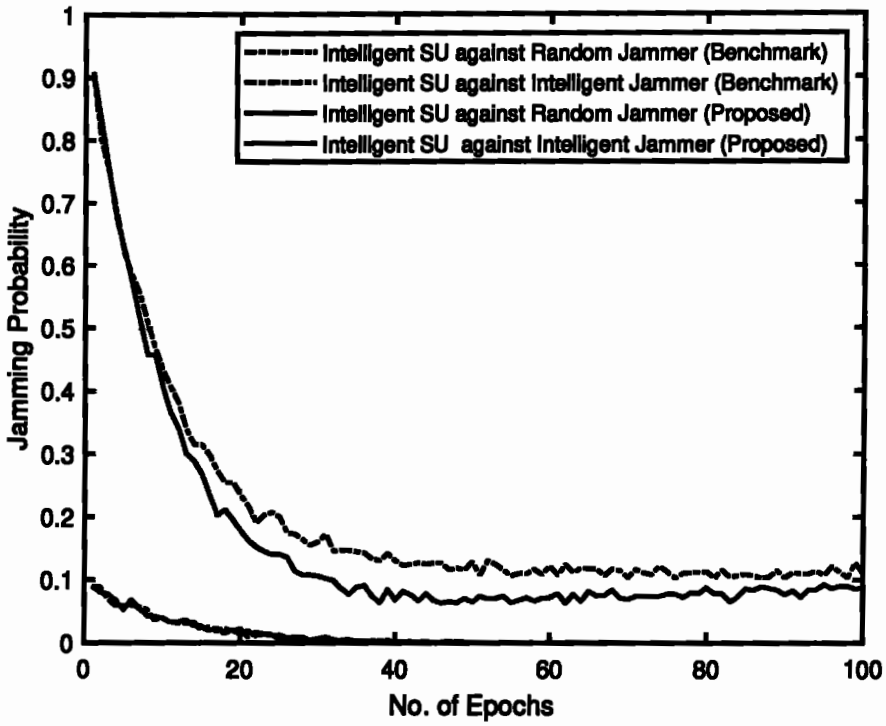
In each epoch the simulations are run $A = 2000$ times to get the average bandwidth efficiency for each SU:

$$\overline{\eta^{n}} = \frac{1}{A}\sum_{a=1}^{A}\eta_{a}^{n}. \tag{4.2.2}$$

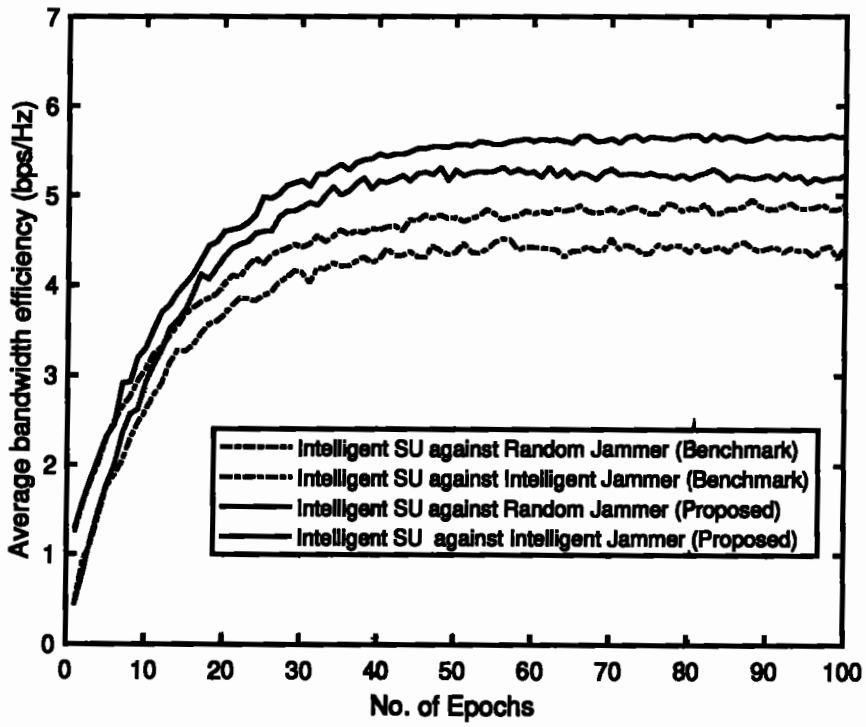
Two cases are considered, where case I refers to the situation when all the $\mathcal{L}$ sub-channels have different SNRs, and hence, different channel capacities as given by Table 4.1. By contrast, all $\mathcal{L}$ sub-channels have the same SNR in case II. More specifically, case II is related to the idealistic scenario [67]. The mean SNR in dB is calculated by:

$$\overline{\mathrm{SNR_{dB}}} = 10\log_{10}\left(\frac{1}{\mathcal{L}}\sum_{l=1}^{\mathcal{L}}10^{\frac{SNR_{dB,l}}{10}}\right), \tag{4.2.3}$$

where $\mathrm{SNR}_{dB,l} = 10\log_{10}(\mathrm{SNR}_{l,t})$ is the SNR of l^{th} sub-channel. The SNR for each sub-channel in case II is the same as the average SNR of case I ($\mathrm{SNR}_{dB} = 16.2$ dB). For a fair comparison, the means SNRs for both cases are equal, which is $\mathrm{SNR}_{dB} = 16.2$ dB as shown in Table 4.1.

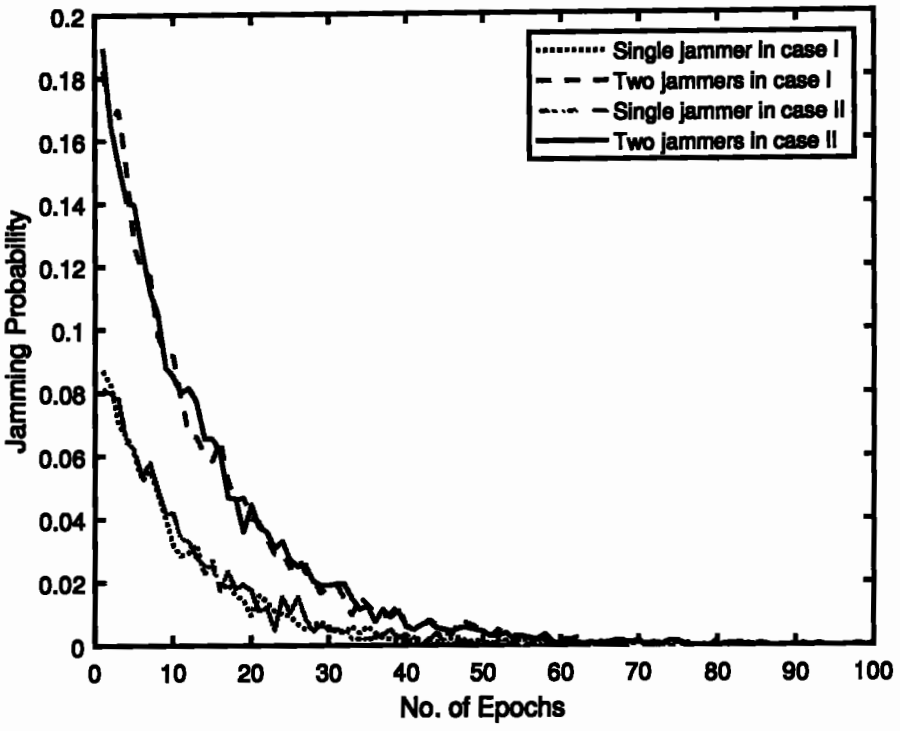


(a) Probability of successful jamming.

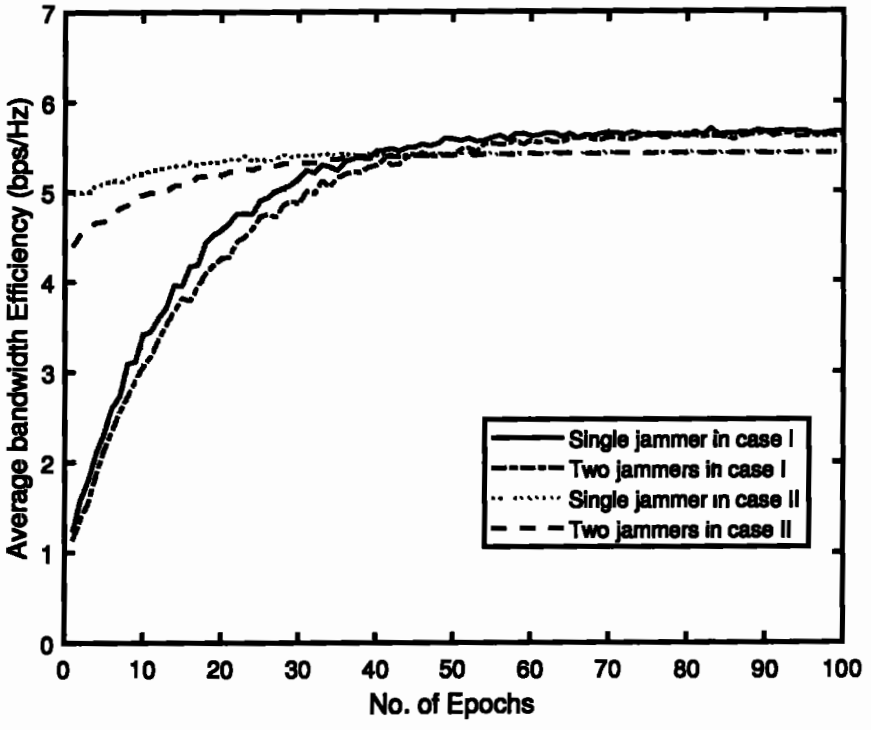


(b) Bandwidth efficiency

Figure 4.1: Comparison of the probability of successful jamming and bandwidth efficiency of the proposed system against benchmark system in case I.



(a) Probability of successful jamming.



(b) Bandwidth efficiency.

Figure 4.2: Comparison of the probability of successful jamming and the bandwidth efficiency for SU, when considering one or two **random jammers** for both cases, case I and case II.

Here, the proposed and benchmark scheme referred in Section 2.2.1 in case I are compared, where the sub-channels have varying qualities. However, the benchmark scheme assumes that all sub-channels have a fixed quality Equation (3.2.6). As seen in Figure 4.1, the proposed system outperforms the benchmark scheme in terms of jamming probabilities and bandwidth efficiency. Please note that when the channel qualities of all sub-channels are fixed, the proposed scheme will perform similarly to the benchmark scheme. In other words, the proposed scheme is the generalization of the benchmark scheme to the general case, where the sub-channel qualities vary.

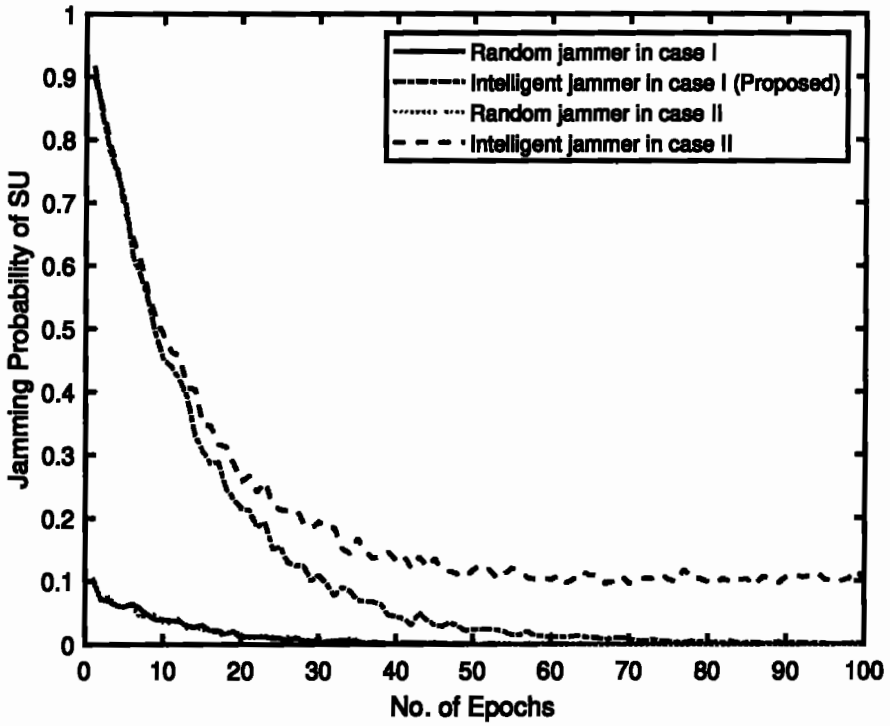
Figure 4.2 shows the probability of successful jamming by the jammer and the bandwidth efficiency of the SU, for both cases I and II, when one or two random jammers are considered. With increasing epochs, the intelligent SU could learn the environment in a better way and the probability of successful jamming is expected to be reduced, while the bandwidth efficiency would increase. The probability of successful jamming for the two-jammer scenario is slightly higher than that of the single-jammer scenario, but the probabilities converge to zero after 60 epochs, as shown in Figure 4.2a. Furthermore, in the more challenging case I, where the channel quality varies across the sub-channels, the proposed algorithm still works well despite requiring a more extended training period (or epochs) to reach

the convergence point, as seen in Figure 4.2b.

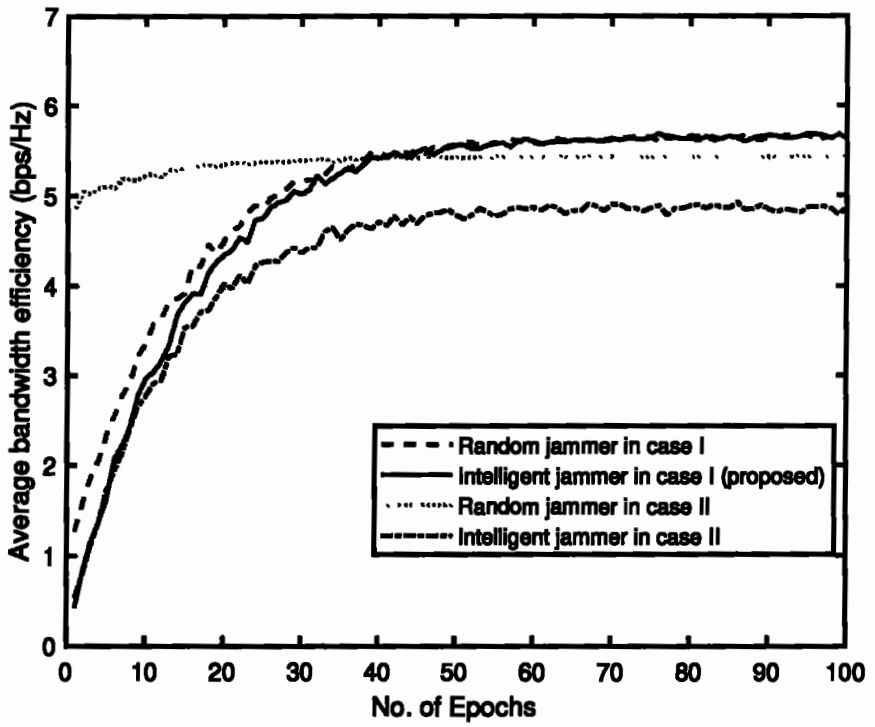
The proposed algorithm allows SU to intelligently choose sub-channels with higher channel capacities when the channels are varying as in the case I. Hence, the average bandwidth efficiency of the system is improved. It can also be seen from Figure 4.2b that the average bandwidth efficiency in the single-jammer scenario of the case I (solid line) is higher than that of the single-jammer scenario of the case II (dotted line) after 40 epochs. A similar pattern can be seen for the two-jammer scenario in Figure 4.2b, after 45 epochs. In other words, the proposed algorithm works better for the case I after a sufficient training period. Hence, the proposed scheme that operates in variable-quality channels (in case II) outperforms the benchmark scheme of [67] that works in fixed-quality channels (in case I). Furthermore, the average bandwidth efficiency of the SU in the two-jammer scenario is almost equal to that of the single-jammer scenario for both case I and case II. This indicates that the proposed algorithm performs equally well when working against two random jammers.

4.2.2 The Effect of Having Different Types of Attacks

In this scenario, the impact of having an intelligent jammer in the system is investigated. Keep in mind that the intelligent jammer also learns from its Q values given in Equation (4.1.8) based on the

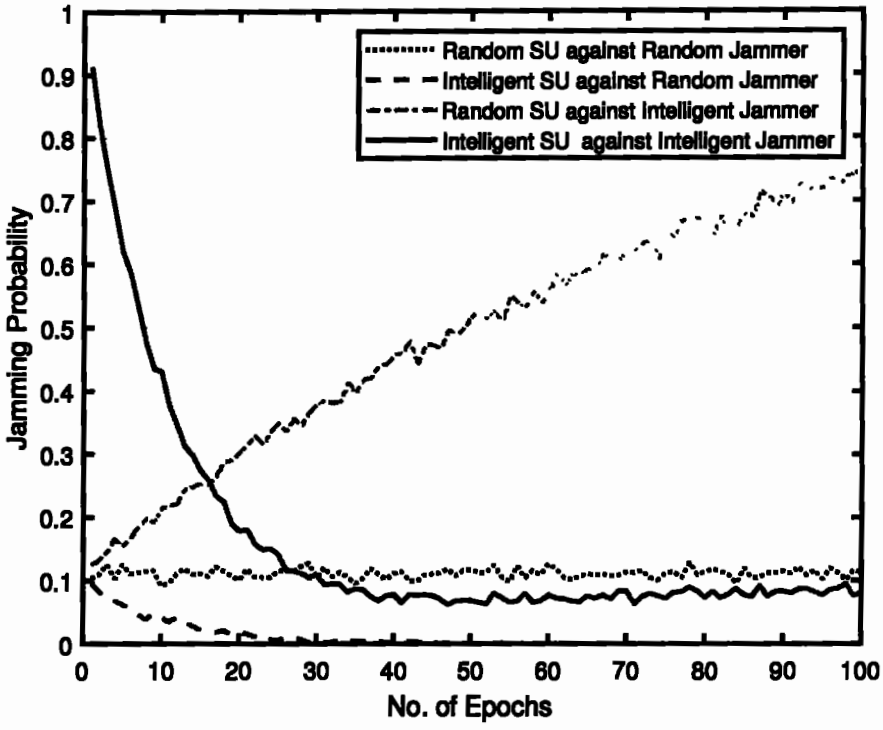


(a) Probability of successful jamming.

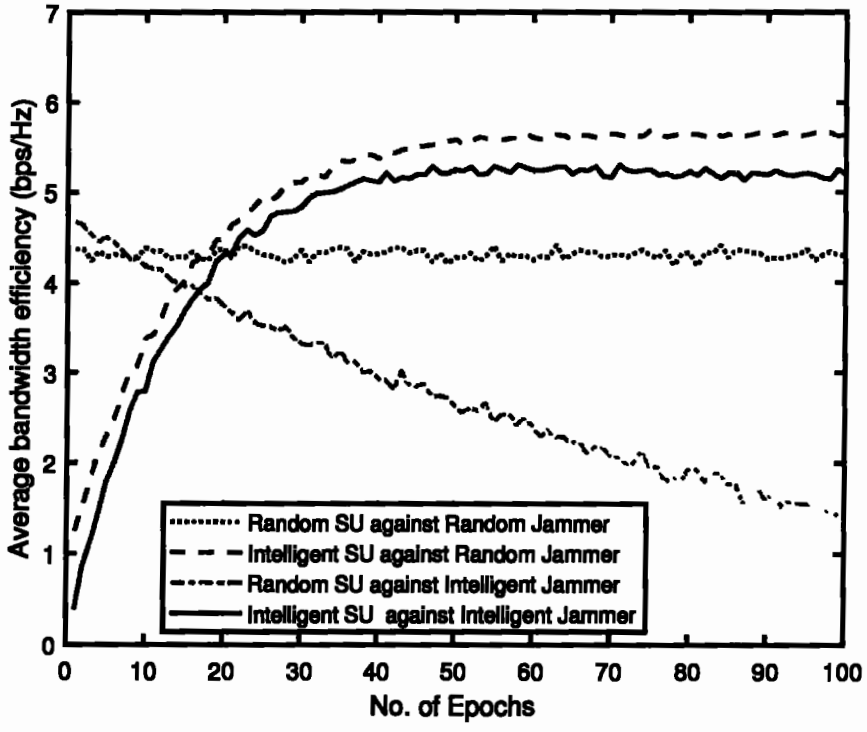


(b) Bandwidth efficiency.

Figure 4.3: Comparison of the probability of successful jamming and bandwidth efficiency for SU, when considering **random** or **intelligent jammers** for both case I and case II. Other parameters are given in Table 4.1.



(a) Probability of successful jamming.



(b) Bandwidth efficiency

Figure 4.4: Probability of successful jamming and bandwidth efficiency of the system having intelligent/random SU against intelligent/random jammer in case I.

parameters given in Table 4.1.

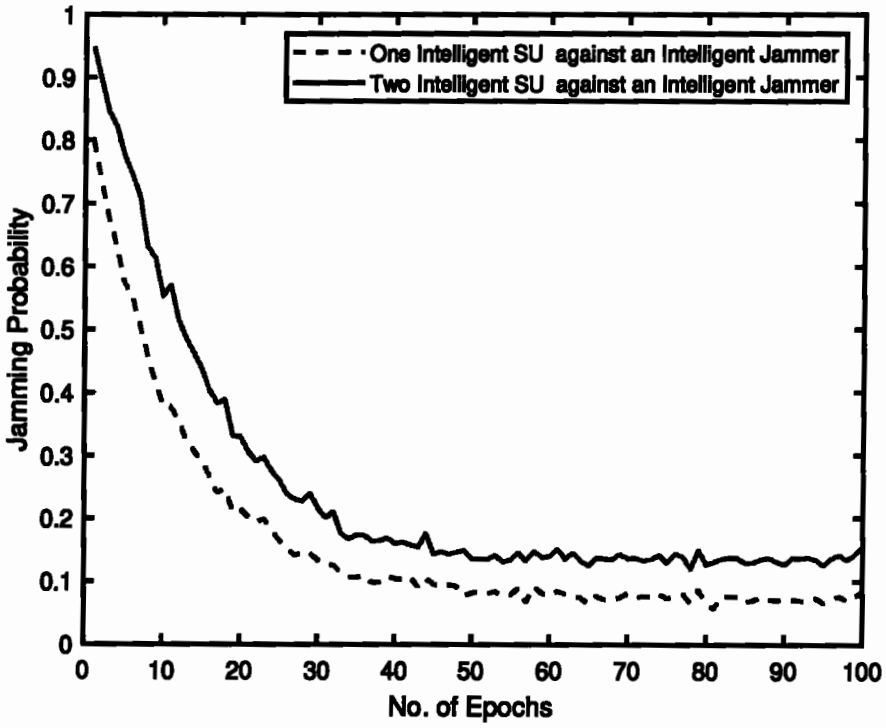
Figure 4.3a shows the jamming probabilities of the random and intelligent jammers for both case I and case II. In particular, the jamming probability converges to zero after 30 epochs when a random jammer is invoked. The successful jamming probability in case I (dashed dot line) is around 10% at the 30^{th} epoch when an intelligent jammer is present. Hence, the successful jamming probability is greater for an intelligent jammer compared to that of a random jammer as expected. The two curves (case I and case II) are almost the same for the random jammer case (not the intelligent jammer case). The proposed scheme (case I) in solid blue is similar to the benchmark scheme (case II) in dotted yellow because of the non-intelligent behavior of the random jammer. From a random jammer perspective, the sub-channel qualities do not matter. Therefore the successful jamming probabilities against a random jammer for case I and case II are almost similar. In contrast, the two curves are different for the intelligent jammer case. The focus of this research is to combat against an intelligent jammer. The intelligent jammer looks for sub-channels with good sub-channel attributes. The proposed scheme works better against intelligent jammer in variable sub-channels case (case I, dash-dot red) as compared to the benchmark scheme (case II, dash indigo). The proposed scheme (case I, dash-dot red) reduces the successful jamming probability to zero af-

ter 70 epochs, while the benchmark scheme (case II, dash indigo) is not capable of doing so.

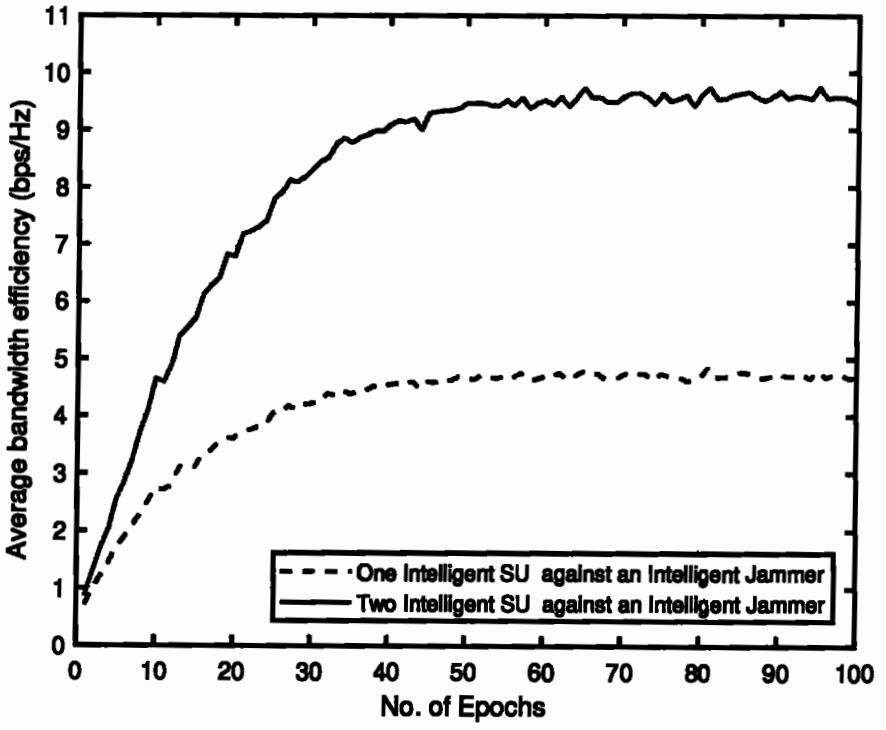
As seen in Figure 4.3b, the corresponding average SU bandwidth efficiency in the presence of an intelligent jammer (solid line) is almost equal to that when having a random jammer (dashed line) after 40 epochs. Hence, the proposed intelligent SU can avoid the intelligent jammer after a certain training period. Furthermore, the SU bandwidth efficiency in case I is higher than in case II. Hence, the SU can also choose intelligently sub-channels with higher capacity, in case I, for increasing the average bandwidth efficiency of the system while successfully avoiding the intelligent jammer.

4.2.3 The Effect of Using Different Defence Strategies

Here, all four possible intelligent/random SU against intelligent/random jammer scenarios are discussed, based on case I. As seen in Figure 4.4, when a SU chooses a random sub-channel strategy in the presence of a random jammer, then both the successful jamming probability and the average bandwidth efficiency of SU remain almost constant (dotted lines). The performance of SU improves remarkably when it behaves intelligently against the random jammer (dashed lines). It is visible that a SU using a random strategy against an intelligent jammer will result in severe jamming and the SU bandwidth efficiency degrades drastically (dashed dotted lines). Hence,



(a) Probability of successful jamming.



(b) Bandwidth efficiency

Figure 4.5: Probability of successful jamming and bandwidth efficiency of the system having multiple intelligent SUs against an intelligent jammer in case I.

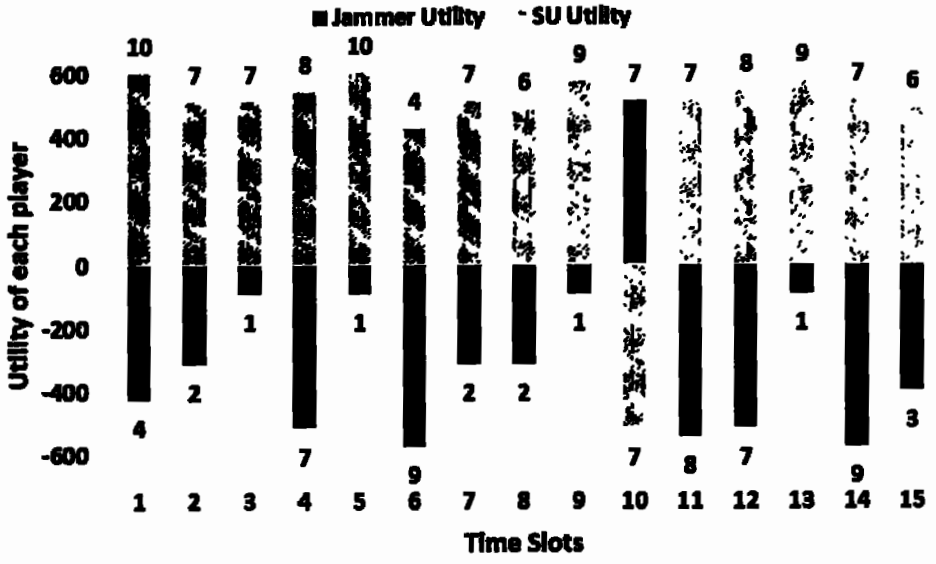


Figure 4.6: Decision pattern in the first 15-time slots for intelligent SU and intelligent jammer.

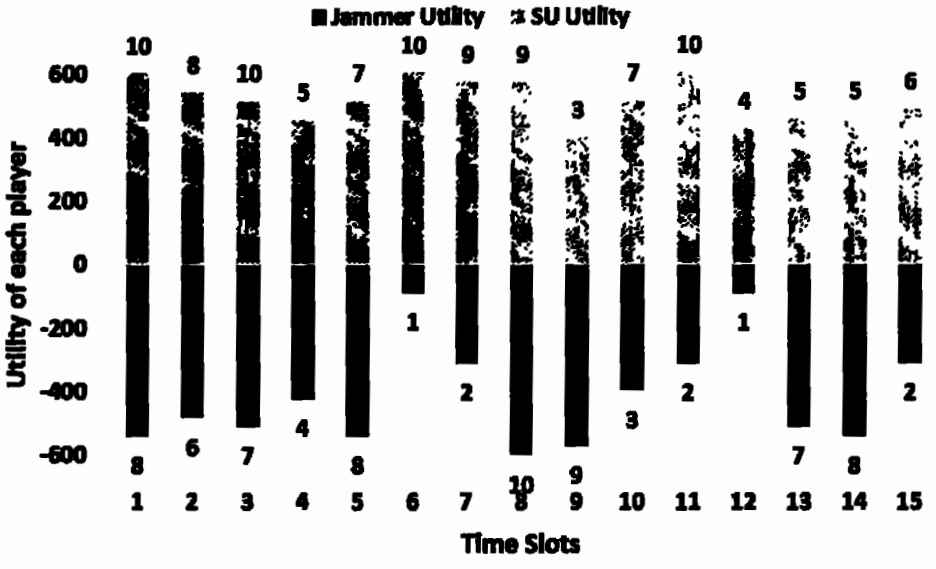


Figure 4.7: Decision pattern in the first 15-time slots for intelligent SU and random jammer. The SU is inclined towards sub-channels with higher SNR.

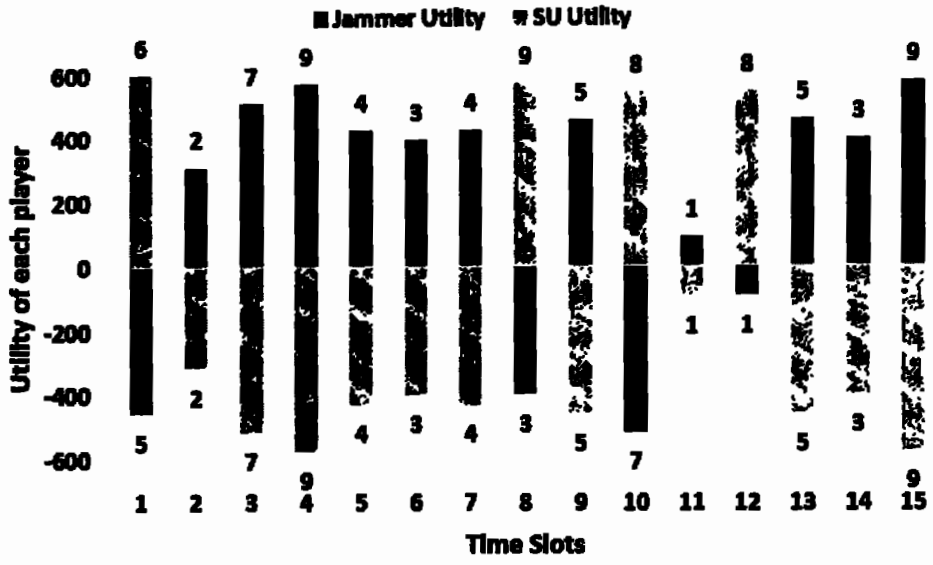


Figure 4.8: Decision pattern in first 15-time slots for random SU and intelligent jammer.

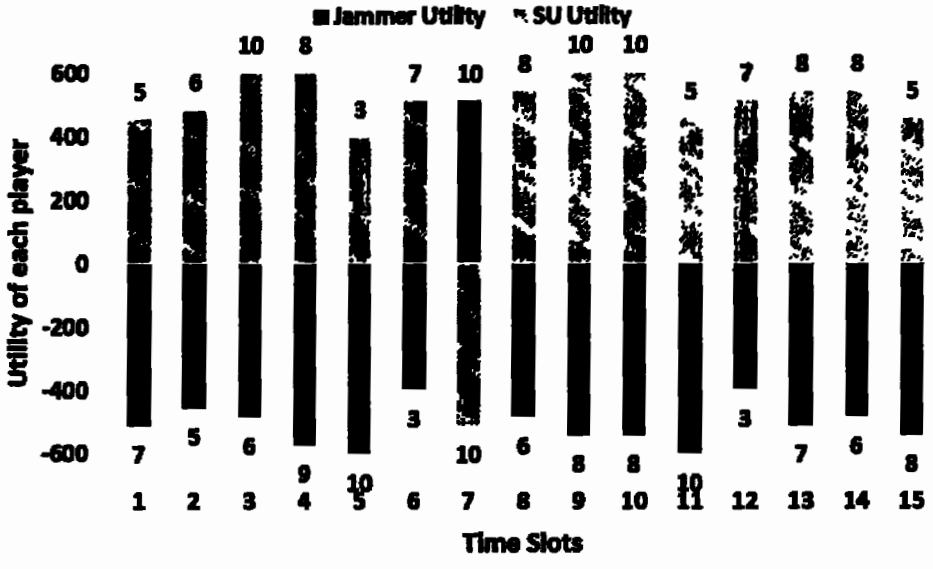


Figure 4.9: Decision pattern in first 15-time slots for intelligent SU and intelligent jammer when the previous action is considered an initial state.

the SU must be intelligent to combat against an intelligent jammer to reduce the successful jamming probability and increase the average SU bandwidth efficiency (solid lines).

4.2.4 The Effect of Multiple Intelligent SUs against Intelligent Jammer

In Figure 4.5 the performances are compared when increasing the number of intelligent SUs in the secondary network. In Figure 4.5a, it is shown that the jamming probability is higher for two intelligent SUs in the presence of a single intelligent jammer compared with the situation when only one intelligent SU is transmitting. The huge impact of an increase in bandwidth efficiency is shown in Figure 4.5b. The bandwidth efficiency of the secondary network is almost doubled when there are two intelligent SUs against an intelligent jammer.

4.2.5 Performance Evaluation of Intelligent SU against Intelligent Jammer in Time Slotted View

In Figure 4.6, Figure 4.7 and Figure 4.8, the x-axis shows the time slot index and the height of the bar shows the utility earned based on the decision of each player after a training period of 100 epochs. As already described in the proposed model, each sub-channel has different quality based on the received SNR. Without loss of generality it is assumed that $\text{SNR}_1 \leq \text{SNR}_2 \leq \cdots \leq \text{SNR}_i \leq \cdots \leq \text{SNR}_{N-1} \leq \text{SNR}_N$, $i \in \mathcal{N}$ i.e., the SNR is increasing from sub-channel 1

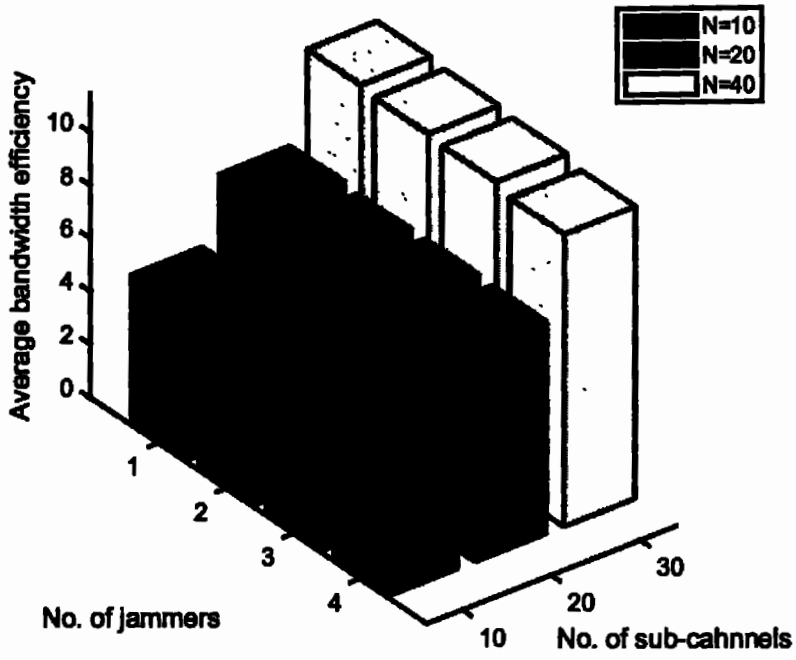


Figure 4.10: The average bandwidth efficiency of an intelligent SU when up-to four random jammers are present in case I, for a different number of available sub-channels.

to sub-channel 10. The impact of changes in SNR and the decision of each player on the utilities are shown in Figure 3.4, where the utilities earned by each player are opposite to each other.

When the players have trained adequately, then the corresponding Q tables would be appropriately updated, which would result in good decisions for all players. The intelligent SU has more choices in terms of choosing an optimal sub-channel. More explicitly, the SU can choose any of the available sub-channels, while there is only one sub-channel that the jammer can choose 'correctly' for successful jamming. Hence, it is less probable for the intelligent SU

to get jammed based on the updated Q table values. Both intelligent SU and intelligent jammer would opt for high-quality channels as depicted by the height of each bar in Figure 4.6. As seen in Figure 4.6, the SU chooses sub-channel 10, while the jammer chooses sub-channel 4 at time slot 1. Hence the SU has a positive utility, while the jammer has a negative utility. Also shown in Figure 4.6, the jammer only manages to jam the SU at time slot 10, over the 15-time slots considered. Hence, the intelligent SU manages to avoid the jammer, while choosing high-capacity sub-channels.

4.2.6 Performance Evaluation of Intelligent SU against Random Jammer in Time Slotted View

Figure 4.7 shows the decision patterns for the case when intelligent SU adapts the Q learning for strategy update, while the jammer uses random strategy. As seen in Figure 4.7, the intelligent SU manages to avoid the random jammer in all of the 15-time slots considered, while at the same time capable of choosing high-capacity sub-channels.

4.2.7 Performance Evaluation of Random SU against Intelligent Jammer in Time Slotted View

Figure 4.8 depicts the decision patterns for the case when the SU uses a random strategy against an intelligent jammer that invokes the Q learning for strategy update. As depicted in Figure 4.8, the jamming

rate is high, which is 11 successful jamming out of the 15-time slots considered.

4.2.8 Performance Evaluation When the Last Action is the Initial State:

Decision pattern in first 15-time slots for intelligent SU and intelligent jammer when the previous action is considered as an initial state in Figure 4.9.

4.2.9 The Effect of Increasing the Number of Sub-channels and Jammers:

Figure 4.10 shows the impact of increasing the number of sub-channels and the number of jammers. As the sub-channel increases, the intelligent SU will have a higher chances of avoiding the jammer. It is shown that the bandwidth efficiency of SU increases as the number of sub-channels increases due to the increase of the choices in the sub-channel space. It is found that increasing the number of random jammers does not significantly affect the bandwidth efficiency of the intelligent SU. It seems that frequency hopping is a very good strategy for the SU to avoid the jammers, especially when the number of sub-channels is high.

4.3 Summary

The major contributions of this chapter are summarized below:

1. For discrete problems like the selection of frequencies in anti-jamming problems, it is difficult to manage using convex optimization [68]. Hence, learning theory is needed in the decision process. The learning algorithm should cope with uncertain dynamics and incomplete information, whereas game theory can adequately model and analyze the mutual interactions among adversarial users. Therefore, it is promising to incorporate the learning algorithm into game theory.

2. Against this backdrop, a game-theoretic optimal frequency hopping scheme is devised between SU and intelligent jammer in a dynamic environment using the Q-learning approach to pick the optimal sub-channel as shown in Figure 3.1.

3. A game model with both players as Independent Learners (ILs) is developed, where they selfishly and independently select their optimal sub-channels in a Multi-Agent Reinforcement Learning (MARL) setting to increase their utilities.

4. The proposed game theoretic model in conjunction with learning based FH algorithm helps SU avoid the attacker and hence reduce the probability of jamming and increase the bandwidth

efficiency of the cognitive system.

5. Moreover, the complexity analysis and the results and discussions are presented to conclude the chapter.

6. The simulation results show improvement in the performance and overall bandwidth efficiency of the cognitive radio network.

CHAPTER 5

Part-II: Enhanced Protection Against a Rate-Aware Intelligent Jammer in Cognitive Radio Networks

This chapter presents the second part of the research work presented in this thesis. This chapter is dedicated to the deception based anti-jamming techniques to combat against a rate-aware jammer in cognitive radio network. As discussed in chapter 3, a PSU is used to deceive the intelligent jammer. The utility of the system is derived for two cases of with and without PSU.

Definition 1[115]: If the following two inequalities are fulfilled, a set of strategies (i^*, j^*) {i for row player, and j for column player} result in a non-cooperative[1] Nash equilibrium solution to a

[1]If the agents in the game receive opposite rewards, then the game is called non-cooperative while if all the agents receive the same reward, the game is called a fully cooperative game.

bimatrix game ($\mathcal{A} = \{A_{ij}\}, \mathcal{B} = \{B_{ij}\}$), where $\mathcal{A}$ and $\mathcal{B}$ are reward matrices for each player, for all $i = 1, 2, 3, ..., N$ and $j = 1, 2, 3, ..., N$ states.

$$\begin{aligned} A_{i^*j^*} &\geq A_{ij^*}, \quad \forall i, j = 1, 2, 3, ..., N \\ B_{i^*j^*} &\geq B_{ij^*}, \quad \forall i, j = 1, 2, 3, ..., N \end{aligned} \tag{5.0.1}$$

Furthermore, the pair $(A_{i^*j^*}, B_{i^*j^*})$ is regarded as the bimatrix game's non-cooperative Nash equilibrium outcome. Where A_{ij} represent the payoff matrix for player I and B_{ij} is the payoff matrix for player II. The payoff matrix of standard matrix game should represent the objective of each player.

5.0.1 Utility without PSU:

The utility function of SU network can be written as the overall gain of the secondary network, which is being controlled by the cognitive base station (CBS) having a number of SUs in the network, is determined by

$$\mathcal{U}_{CBS} = \sum_{i=1}^{N} \mathcal{U}_i \tag{5.0.2}$$

where $\mathcal{U}_i = \mathcal{C}_i((1 - \delta_i)\mathcal{G} - \delta_i\mathcal{J}), \quad \forall 1 \leq i \leq N$ and $\mathcal{C}_i = \mathcal{W}/\mathcal{L}\log_2(1 + \mathcal{P}_{r,i}/(\mathcal{N}_w.\mathcal{W}/\mathcal{L})), \quad \forall 1 \leq i \leq N$ which satisfy

$$\begin{cases} \delta_i = \mathcal{R}_i/\sum_{j=1}^{N} \mathcal{R}_j, \quad \forall i \in \mathcal{N} \\ \sum_{i=1}^{N} \delta_i = 1 \\ 0 \leq \delta_i \leq 1 \end{cases} \tag{5.0.3}$$

where $\mathcal{R}_i$ is the data rate based on the modulation schemes adapted by the i^{th} SU/PSUs. The intelligent jammer may acquire the rate/code/modulati

information of SU/PSU either using explicit rate information or modulation and code guessing [52]. Moreover, the rate information of a transmission is vulnerable in many communication protocols. In IEEE 802.11 networks, for example, the rate is specified explicitly in the SIGNAL field of the physical layer's frames. An intruder can easily coordinate with two parties' communication, evaluate data frames, and derive the rate. As demonstrated in [49], this attack is quite practical. The adversary can evaluate the received signal in complicated I/Q form even if the rate information is not explicitly supplied inside the packet header. The attacker can trace the received constellation pattern and determine the modulation in use after performing carrier synchronization, frequency, and phase offset correction. The frame structure of the protocol is not required for this method. The guessing strategy on USRP can be shown by creating a modulation detector that can identify the modulation of a transmission in real time. It may readily be modified to create a practical rate-aware jammer that jams high-rate packets selectively. An attacker could employ more sophisticated techniques to determine not only the modulation of the message, but also the codes used. One such method is to follow the sequence of received symbols in order to predict the codes based on the fact that various codes cause distinct transitions from one coded symbol to the next. For the attacker, guessing through matching and trial-and-error is effi-

cient since most communication protocols specify a finite variety of modulations and codes [52].

5.0.2 Utility with PSU:

The utility function of secondary network in the presence of PSU can be written as the overall gain of the secondary network, which is being controlled by the cognitive base station (CBS) having few SUs and a PSU, is determined by $\mathcal{U}'_{CBS} = \sum_{i=1}^{N-1} \mathcal{U}'_i + \mathcal{U}'_{PSU}$. Since the PSU does not take part in useful communication, so the utility of PSU is not counted for the calculations of throughput. Therefore, the ultimate utility of CBS is given by

$$\mathcal{U}'_{CBS} = \sum_{i=1}^{N-1} \mathcal{U}'_i, \tag{5.0.4}$$

where $\mathcal{U}'_i = C_i((1-\delta'_i)\mathcal{G} - \delta'_i\mathcal{J}), \quad \forall 1 \leq i \leq N$, and satisfies,

$$\begin{cases} \delta'_i = \mathcal{R}_i/(\sum_{j=1}^{N-1}\mathcal{R}_j + \mathcal{R}_p), \quad \forall i \in \mathcal{N} \\ \delta'_p = \mathcal{R}_p/(\sum_{j=1}^{N-1}\mathcal{R}_j + \mathcal{R}_p), \\ \sum_{i=1}^{N-1}\delta'_i + \delta'_p = 1 \\ 0 \leq \delta'_i, \delta'_p \leq 1 \end{cases} \tag{5.0.5}$$

where δ'_i is the attraction factor for i_{th} user in the presence of PSU in the network and $\mathcal{R}_p$ is the data rate of deceptive PSU transmission. The attraction factor δ_p of a PSU is kept slightly higher than all other legitimate SUs in the network. i.e., $\delta_i^i < \delta_p$, so that the jammer is more attracted towards PSU as compared to legitimate SU. In general, the probability of the attacker falling in the trap of PSU is

given by $\delta_p = 1 - \sum_{i=1}^{N-1} \delta_i'$.

The objective of the secondary network is to maximize the network utility by successfully deploying PSU to deceive the jammer while increasing the throughput of the system. The throughput of the system is calculated by adding the throughput of every successful individual user, i.e.,

$$U_{CBS}' = U_1' + U_2' + U_3' + \cdots + U_{N-1}' \tag{5.0.6}$$

Using the values from Equation (3.3.3) and rearranging will give

$$\begin{aligned} U_{CBS}' = \mathcal{G} \sum_{i=1}^{N-1} C_i - (\mathcal{G}+\mathcal{J})\delta_1' C_1 + (\mathcal{G}+\mathcal{J})\delta_2' C_2 \\ +(\mathcal{G}+\mathcal{J})\delta_3' C_3 + \cdots + (\mathcal{G}+\mathcal{J})\delta_{N-1}' C_{N-1} \end{aligned} \tag{5.0.7}$$

which further reduces to

$$U_{CBS}' = \mathcal{G} \sum_{i=1}^{N-1} C_i - (\mathcal{G}+\mathcal{J}) \sum_{i=1}^{N-1} \delta_i' C_i. \tag{5.0.8}$$

The problem can be formulated as an optimization problem. The optimal strategy of the CBS is the maximizer of the following problem

$$\max U_{CBS}' = \mathcal{G} \sum_{i=1}^{N-1} C_i - (\mathcal{G}+\mathcal{J}) \sum_{i=1}^{N-1} \delta_i' C_i. \tag{5.0.9}$$

subject to

$$\begin{cases} \delta_i' = \mathcal{R}_i / (\sum_{i=1}^{N-1} \mathcal{R}_i + \mathcal{R}_p), & \forall i \in \mathcal{N} \\ \delta_p' = \mathcal{R}_p / (\sum_{i=1}^{N-1} \mathcal{R}_i + \mathcal{R}_p), & \forall i \in \mathcal{N} \\ \sum_{i=1}^{N-1} \delta_i' + \delta_p' = 1 \\ 0 \le \delta_i', \delta_p' \le 1 \end{cases} \tag{5.0.10}$$

5.0.3 Cost of Implementing PSU-based Deception

The cost is paid in terms of bandwidth loss incurred by adapting a dummy user named as a pseudo secondary user (PSU). The loss is higher when the PSU occupies a sub-channel with higher SNR values, resulting in more throughput reduction as depicted in Figure 5.6 Figure 5.5 and the discussion hereafter. Since the PSU does not take part in the bandwidth efficiency of the CBS, therefore, as the PSU hops to the higher quality sub-channel, the data rate that could be utilised otherwise is wasted by adapting the PSU to deceive the jammer.

5.1 Results and Discussions

The simulation results are shown in bandwidth efficiency. The bandwidth efficiency can be calculated in bits per second per Hertz (bps/Hz) as follows:

$$\eta^n = \frac{C^n_{l,t}(a^n, c^m_j)}{\frac{\mathcal{W}}{\mathcal{L}}} = \log_2\left(1 + \frac{\mathcal{P}^n_{l,t}}{\mathcal{N}_o.\mathcal{W}/\mathcal{L}}\right), \quad \forall n \in \mathcal{N} \tag{5.1.1}$$

The proposed results are compared with and without the PSU in the network. i.e., the comparison is made for N SUs against $N-1$ SUs plus 1 PSU. for simplicity, the channel conditions are considered in ascending order i.e., $SNR_1 < SNR_2 < SNR_3 < \cdots < SNR_l < \cdots < SNR_L$. It means that the L^{th} channel is the best channel among all.

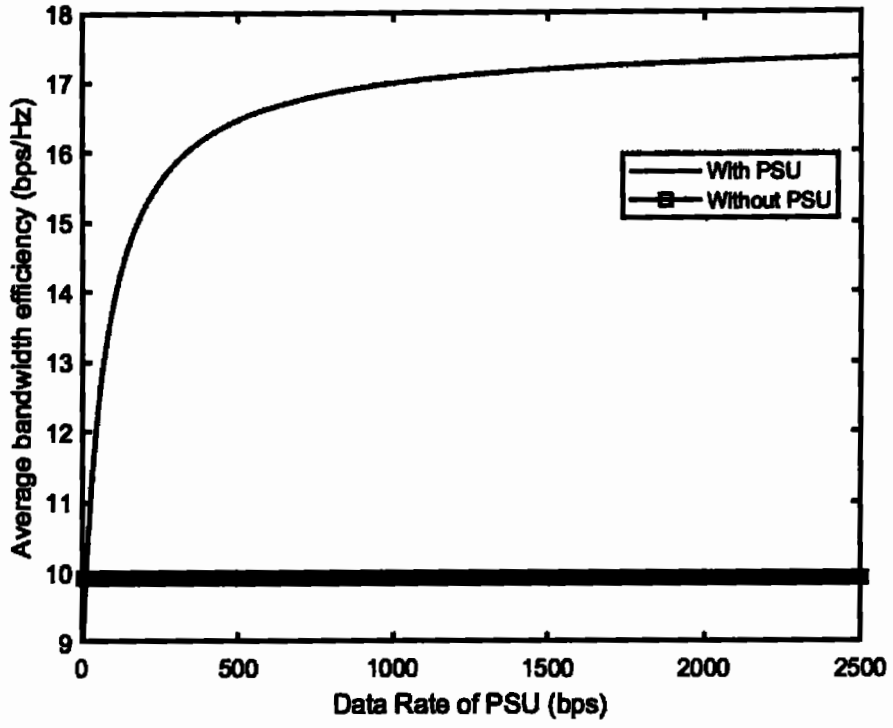


Figure 5.1: Impact of increasing PSU data rate against a rate ware jammer.

5.1.1 The Impact of the Increasing Data Rate of PSU against a Rate-aware Intelligent Jammer

Figure 5.1 depicts the impact of the increasing data rate of PSU against a rate-aware intelligent jammer. It is shown that increasing the data rate of a PSU increases its attraction factor and become a more attractive target for an intelligent jammer. Hence protects the other SUs from being jammed. The successful communication of rest of $N-1$ SUs increase the overall utility of the system. Specifically, it is shown that the utility almost remains constant after the data rate exceeds 1000 bps. Moreover, the data rate of PSU should be at least great than 15 bps to get higher utility as compared to the

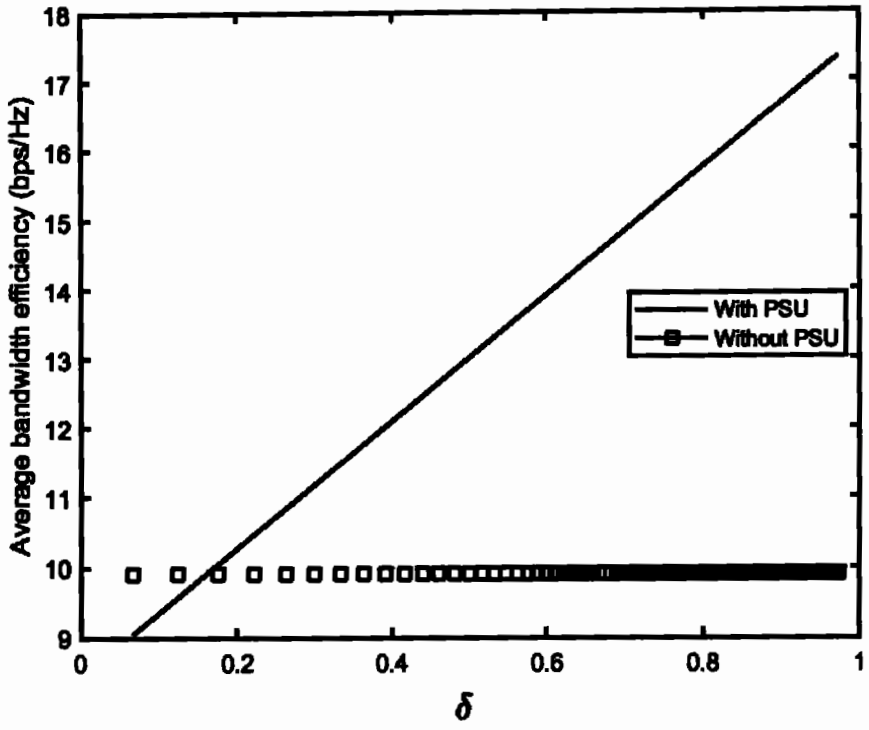


Figure 5.2: Impact of increasing PSU data rate and attraction factor against a rate ware jammer.

system without PSU.

5.1.2 The Impact of Increasing Jamming Probabilities of PSU against a Rate-aware Intelligent Jammer

In Figure 5.2 impact of increasing jamming probabilities of PSU is studied against a rate-aware intelligent jammer. It is shown that increasing the jamming probability of PSU will secure the overall system by increasing the overall utility of the CBS. More precisely, it is evident that the jamming probability of PSU should be at least 0.18 to benefit from deploying PSU in the system. It should be kept in mind that the curve with PSU is considered for $N-1$ users. The

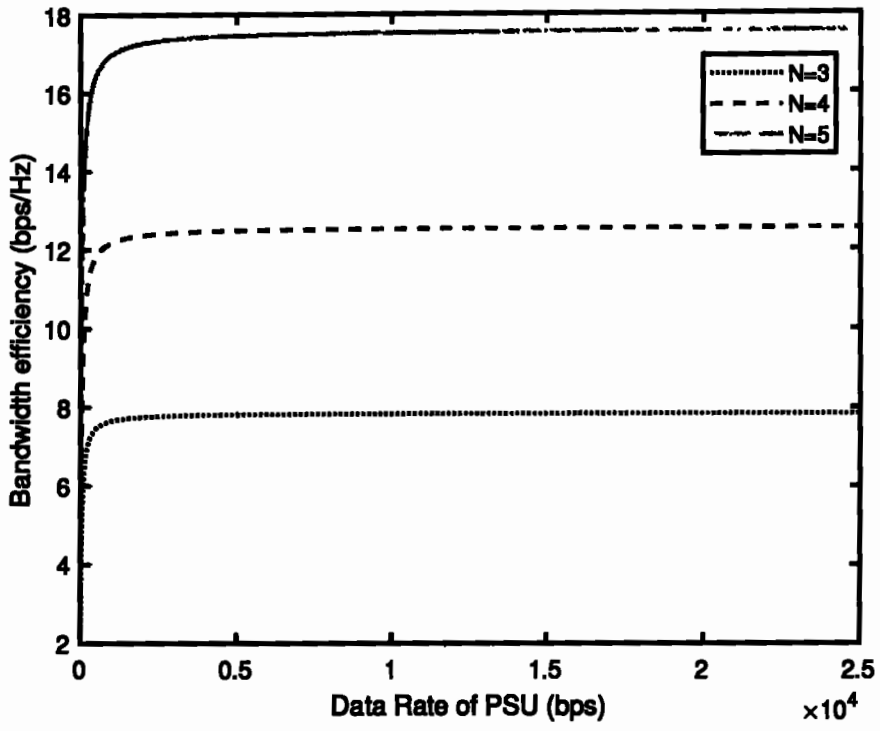


Figure 5.3: Comparison of the data rate of PSU in terms of average bandwidth efficiency for different values of N.

N_{th} user excluded being PSU itself, because it does not transmit legitimate signal and does not take part in overall throughput of the system as mentioned in Equation (5.0.4). It is to be noted that the straight line of the utility of the CBS without PSU (blue rectangles) is due to the fact that it is independent of PSU probabilities. If equal resource utilization for SU/PSU is assumed, then the network with one SU and one PSU consume 50% of its resources to protect one SU. Similarly, when there are seven SUs and one PSU, the $1/8 = 12.5\%$ resources are wasted. It is better to use PSU for a bigger network of SUs.

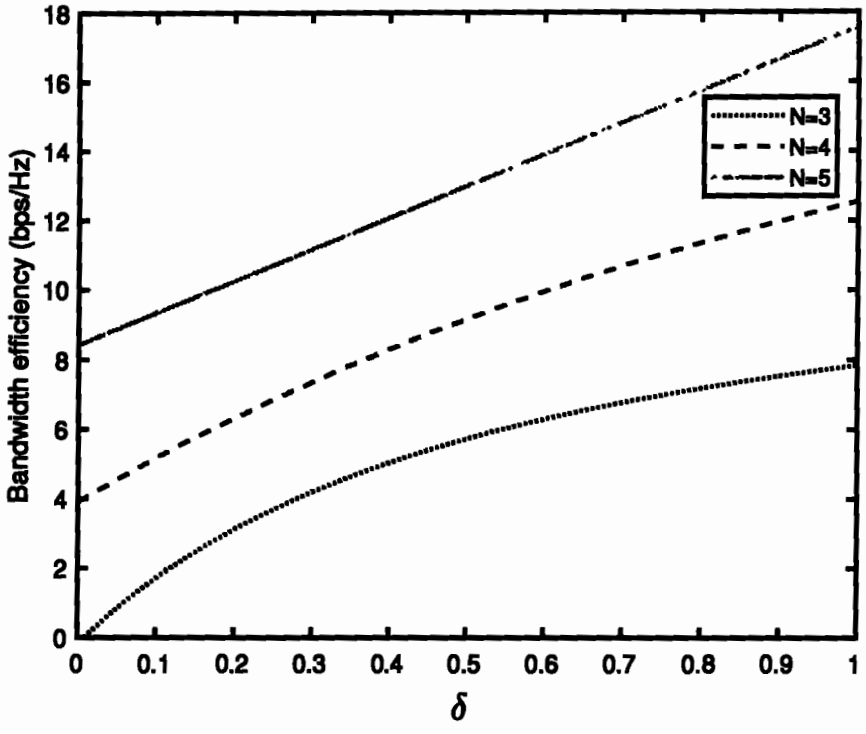


Figure 5.4: Comparison of attraction factor δ of PSU in terms of average bandwidth efficiency for different values of N.

5.1.3 Performance Evaluation for Different Values of N

In Figure 5.3 the curves for different values of N are shown. It has been demonstrated that when N increases, bandwidth efficiency improves. Similarly, the curves in Figure 5.4 show the improvement in bandwidth efficiency as N increases from 3 to 5 against the attraction factor of PSU. A similar behaviour can be observed for N greater than 5.

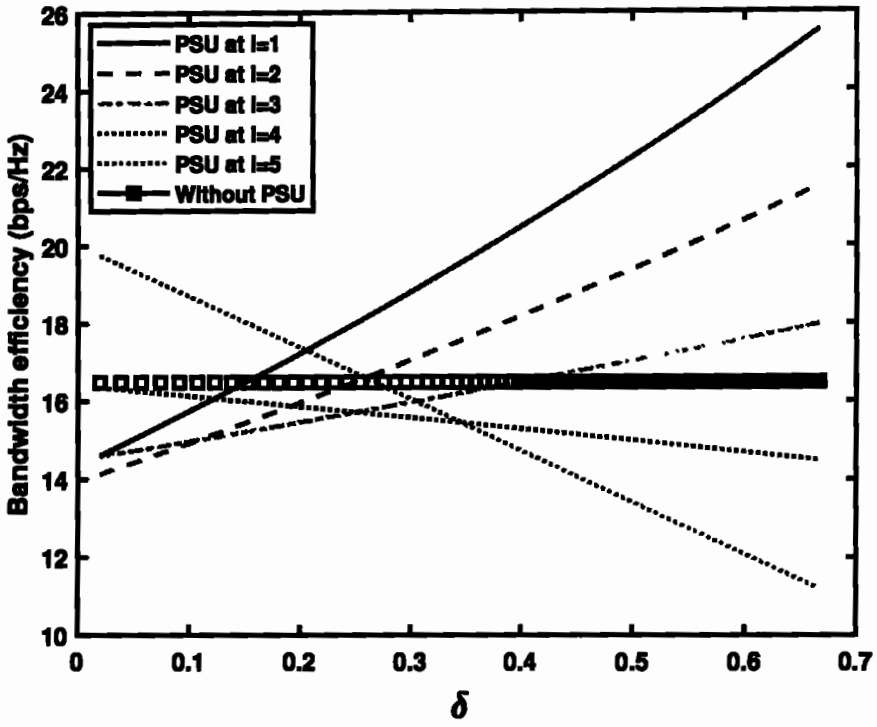


Figure 5.5: Comparison of the data rate of PSU in terms of bandwidth efficiency for different locations of PSU from 1 to L.

5.1.4 The Effect of Changing PSU Positions

The change in the position of PSU is portrayed in Figures 5.5 and 5.6. Figure 5.5 shows the curves for bandwidth efficiency when the position of PSU is changed from 1 to L^{th} sub-channel [2]. The overall bandwidth efficiency of the CBS is highest when PSU is deployed at the sub-channel with the lowest channel quality, as shown by the solid line for $l = 1$. Furthermore, from Figure 5.6 it is evident that for $l = 1$ the PSU is effective only after the attraction factor exceeds 0.15 probability. However, from both Figures 5.5 and 5.6 it

[2]For simplicity, $L = N$. Keep in mind the assumption that the SNR_1 of sub-channel 1 is minimum while SNR_L is maximum.

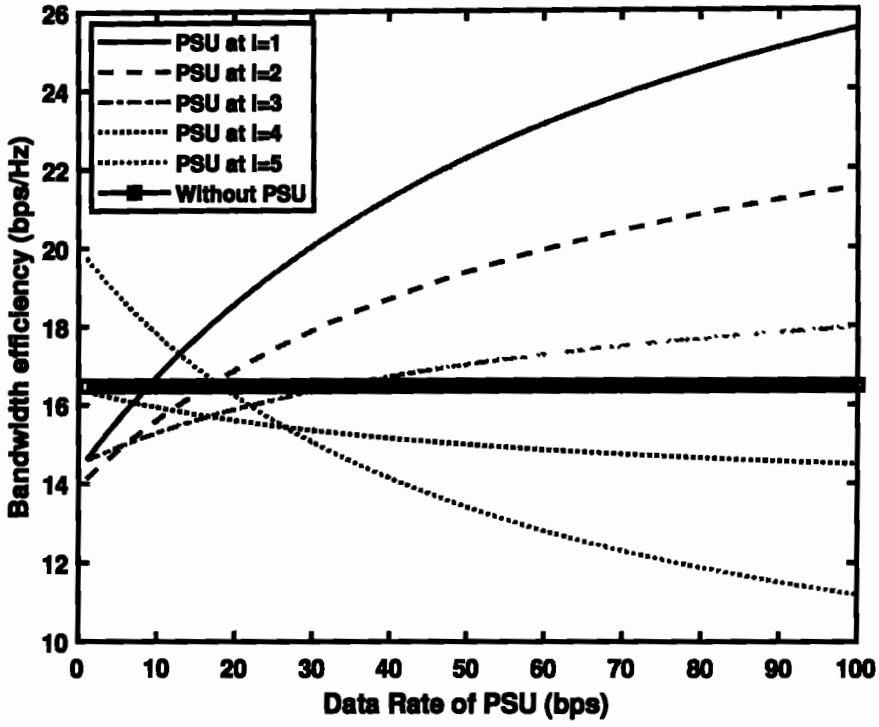


Figure 5.6: Comparison of attraction factor δ of PSU in terms of bandwidth efficiency locations of PSU from 1 to L.

is visible that when the PSU moves to the next sub-channel with higher channel qualities, it may waste its resources to deceive the jammer resulting in less bandwidth efficiency. For a total of $L = 5$ sub-channels, when PSU is at $l = 3$, the bandwidth efficiency curve of PSU crosses the curve without PSU at around 0.43 probability with the lowest productive bandwidth efficiency. The PSU at $l = 4$ and $l = 5$, the bandwidth efficiency gets even worst, which decreases with increase in data rate.

The proposed model results are compared with [1] and [2] to show the bandwidth efficiency of the proposed approach as shown in

Table 5.1. Since the work presented in [1] and [2] use FH and FH + RA, respectively, using reinforcement learning to maximize bandwidth efficiency of the system, therefore, for the fair analysis, the results are tabulated for the better comparison of the three papers.

Table 5.1: Comparison of performance evaluations of our proposed scheme with the work presented in [1] and [2].

Parameters	[1]	[2]	**The proposed scheme**
H: No. of PU	1	1	1
N: No of SU	1	1	1
M: No. of jammers	1	1	1
L: No. of sub-channels	10	5	5
SNR_{dB}	[0 10:20]	-	[0 10:20]
Jammer type	Level-I	Level-I	Level-II
Bandwidth efficiency (bps /Hz)	5.7	4.2	17.24

As depicted in Table 5.1, all other parameters are the same except for jammer type. The jammer type used in [1] and [2] is a reactive jammer, which commences its transmission upon detecting SU activity in the cognitive radio enabled network. This type of jammer is named as Level-I jammer in section 3.3.1. In the proposed work, a more challenging jammer is considered, which is a reactive jammer and targets the SU based on each user's attraction factor, making it more harmful to the highest impact communication.

Moreover, the bandwidth efficiency comparison results demonstrate that employing the deception-based anti-jamming technique proposed in this study has an advantage over the work in [1]. The bandwidth efficiency of the proposed approach is more than 3 times that of [1]. The bandwidth efficiency of the proposed approach is

more than 4 times that of [2]. To be very accurate, the results in [2] shows the bandwidth of 21bps for 5 sub-channels (see Figure 7 in [2]). To get bandwidth efficiency for each sub-channel 21 is divided by 5 to get the bandwidth efficiency equal to 4.2 bps/Hz. Furthermore, all results were obtained with the help of the MATLAB simulation environment.

5.2 Summary

The major contributions of this chapter are summarized below.

1. The focus of this chapter was on combating smart rate-aware jamming attacks by adjusting transmission rate in cognitive radio networks. An intelligent reactive jammer was considered, which was called an infant jammer in section 3.3.1.

2. A cognitive adversarial rate aware jammer, which is an intelligent attacker aware of the communication parameters i.e., transmission rate and can adapt the dynamics of the sub-channels and the strategies of SU.

3. A unique utility function is introduced, where the channel conditions may change from one sub-channel to another with near practical channel conditions.

4. The mathematical model of a novel deception-strategic pseudo secondary user (PSU) is proposed by introducing an attraction

factor δ of each user based on the actual transmission rate to decoy intelligent rate aware jammer.

5. The simulation results show improvement in the performance and overall bandwidth efficiency of the cognitive radio network.

CHAPTER 6

Conclusions and Future Suggestions

Conclusion and the future dimensions in the light of limitations are presented subsequently.

6.1 Conclusions

This research investigates an anti-jamming stochastic game in conjunction with a multi-agent reinforcement learning algorithm. Both random and intelligent jammers were considered. The anti-jamming game was designed as a Markov game based on the Q-learning algorithm. A game-theoretic optimal frequency hopping scheme in a dynamic environment is devised in the presence of adversarial jammers by using Q-learning approach to pick high-capacity subchannels while avoiding the jammer. A game model with both players as independent learners is developed, where SU and jam-

mer selfishly and independently select their optimal sub-channels in a multi-agent reinforcement learning setting to increase their individual utilities. The proposed game theoretic model in conjunction with learning based frequency hopping algorithm, helps the SU to avoid the attacker, hence reducing the probability of jamming and increasing bandwidth efficiency of SU. It was shown in the simulation results that the proposed method outperforms the benchmark system in terms of both the bandwidth utilization and the jamming probability. More specifically, the average bandwidth efficiency improves from 4.9 (bps/Hz) to almost 5.7 (bps/Hz) as compared to the benchmark scheme. While the jamming probability is reduced to less than 0.1 using the proposed approach. Furthermore, when the channel exhibits variable channel quality (as in case I), the intelligent SU can intelligently choose sub-channels with higher capacity while avoiding the intelligent jammer. Moreover, the bandwidth efficiency of the SU does not decrease significantly when the number of random jammers increases.

In the **second part of the study**, a unique game-theoretic anti-jamming deception strategy is introduced to improve the overall bandwidth efficiency of a cognitive radio-based communication system. A defensive deception anti-jamming method based on rate modifications is used to deceive the attacker and safeguard the remainder of the network from adversarial effects. To lure the jammer,

a counterfeit user inside the network is introduced as a trap for adversary. The higher a secondary user's data rate is, the more attractive to a rate-aware intelligent jammer. The simulation findings suggest that utilising the recommended deception-based technique has improved performance significantly. Simulation results show that bandwidth efficiency of the network adapting the proposed deception strategy crosses the bandwidth efficiency curve of the network without PSU at around the attraction factor of 0.16, which conform to the claim that the CBS with the PSU performs well even with the attraction factor of 0.20 compared to the system without deception strategy. Furthermore, compared to a system that does not use the deception method, the proposed solution can increase bandwidth efficiency by up to 1.7 times. Similarly, since the PSU does not take part in legitimate communication, assigning PSU the highest quality sub-channel will reduce the bandwidth efficiency hence demanding an optimal sub-channel selection for better bandwidth efficiency, evident from the results shown.

6.2 Future Suggestions

The research presented in this thesis can be extended further in the light of limitations as follow:

1. The proposed anti-jamming research is done with the assump-

tion of perfect time and frequency synchronization. The same can be extended for imperfections in time and frequency synchronizations, in the presence of a jammer.

2. Keeping the role constant of SU can make the jammer conscious about the PSU, and can easily result in the counter-deception strategy by the intelligent jammer. The deception strategy can be extended further if the role of a SU is changed dynamically.

3. Inspired by the **change of guard ceremony** with transitions between SU and PSU may lead to further confusion for the jammer.

4. A dynamic assignment of PSU based on the current dynamics of the environment for the optimized selection of PSU can lead to more complexity in the deception mechanism.

5. Moreover, a PSU with a higher attraction factor can also be used to detect the jammer and predict the intelligence level by guessing its fingerprints.

6. Finally, Using FH with RA for the decoy mechanism may enhance the deception further. Enabling PSUs to hop to other available sub-channels and rate adaption can give the PSU freedom of efficient spectrum utilization while deceiving jammer.

Bibliography

[1] K. Ibrahim, I. M. Qureshi, A. N. Malik, and S. X. Ng, "Bandwidth-efficient frequency hopping based anti-jamming game for cognitive radio assisted wireless sensor networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–5.

[2] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless systems," *IEEE Transactions on Mobile Computing*, vol. 15, no. 9, pp. 2247–2259, 2015.

[3] F. C. Commission *et al.*, "Unlicensed operation in the tv broadcast bands and additional spectrum for unlicensed devices below 900 mhz in the 3ghz band," *ET Docket*, no. 04-186, 2004.

[4] "Pakistan Table of Frequency Allocations Pakistan Telecommunication Authority Frequency Allocation Board Government of Pakistan," *Pakistan Telecommunication Authority.*, no. Pakistan Table of Frequency Allocation, pp. 3–4. [Online].

Available: http://www.pta.gov.pk/media/Pakistan_Table_of_Frequency_Allocations.pdf

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

[6] J. Mitola, "Cognitive radio architecture evolution," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 626–641, 2009.

[7] L. Gavrilovska, V. Atanasovski, I. Macaluso, and L. A. DaSilva, "Learning and reasoning in cognitive radio networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1761–1777, 2013.

[8] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE journal on selected areas in communications*, vol. 23, no. 2, pp. 201–220, 2005.

[9] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A Survey on Machine-Learning Techniques in cognitive Radios," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1136–1159, 2013.

[10] F. Akhtar, M. H. Rehmani, and M. Reisslein, "White space: Definitional perspectives and their role in exploiting spectrum opportunities," *Telecommunications Policy*, vol. 40, no. 4, pp. 319–331, 2016.

[11] S. Chen, K. Zeng, and P. Mohapatra, "Hearing is believing: Detecting mobile primary user emulation attack in white space," in *2011 Proceedings IEEE INFOCOM.* IEEE, 2011, pp. 36–40.

[12] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE transactions on vehicular technology*, vol. 60, no. 7, pp. 3386–3407, 2011.

[13] Y. Saleem and M. H. Rehmani, "Primary radio user activity models for cognitive radio networks: A survey," *Journal of Network and Computer Applications*, vol. 43, pp. 1–16, 2014.

[14] K. Grover, A. Lim, and Q. Yang, "Jamming and anti–jamming techniques in wireless networks: a survey," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 17, no. 4, pp. 197–215, 2014.

[15] H. Tran, G. Kaddoum, F. Gagnon, and L. Sibomana, "Cognitive radio network with secrecy and interference constraints," *Physical Communication*, vol. 22, pp. 32–41, 2017.

[16] S. Parvin, F. K. Hussain, O. K. Hussain, S. Han, B. Tian, and E. Chang, "Cognitive radio network security: A survey," *Journal of Network and Computer Applications*, vol. 35, no. 6. pp. 1691–1708, 2012.

[17] S. Sodagari, A. Attar, V. C. Leung, and S. G. Bilén, "Denial of service attacks in cognitive radio networks through channel eviction triggering," in *2010 IEEE Global Telecommunications Conference (GLOBECOM) 2010.* IEEE, 2010, pp. 1–5.

[18] K. Pelechrinis, M. Iliofotou, and S. V. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Communications surveys & tutorials*, vol. 13, no. 2, pp. 245–257, 2010.

[19] S. Anand, Z. Jin, and K. Subbalakshmi, "An analytical model for primary user emulation attacks in cognitive radio networks," in *2008 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks.* IEEE, 2008, pp. 1–6.

[20] H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems, Part I: Known Channel Statistics," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3566–3577, nov 2010.

[21] K. Li and J. Wang, "Optimal joining strategies in cognitive radio networks under primary user emulation attacks," *IEEE Access*, vol. 7, pp. 183 812–183 822, 2019.

[22] A. Ahmadfard, A. Jamshidi, and A. Keshavarz-Haddad, "Probabilistic spectrum sensing data falsification attack in cog-

nitive radio networks," *Signal Processing*, vol. 137, pp. 1–9, 2017.

[23] B. Rashid, M. H. Rehmani, and A. Ahmad, "Broadcasting strategies for cognitive radio networks: Taxonomy, issues, and open challenges," *Computers and Electrical Engineering*, vol. 52, pp. 349–361, 2016.

[24] 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, aug 2013.

[25] Z. Sun, Y. Liu, J. Wang, G. Li, C. Anil, K. Li, X. Guo, G. Sun, D. Tian, and D. Cao, "Applications of game theory in vehicular networks: A survey," *IEEE Communications Surveys & Tutorials*, 2021.

[26] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.

[27] J. Hu and M. P. Wellman, "Multiagent reinforcement learning: theoretical framework and an algorithm." in *ICML*, vol. 98. Citeseer, 1998, pp. 242–250.

[28] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for distributed dynamic spectrum access," *IEEE*

Transactions on Wireless Communications, vol. 18, no. 1, pp. 310–323, 2018.

[29] B. Wang, Y. Wu, K. R. Liu, and T. C. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE journal on selected areas in communications*, vol. 29, no. 4, pp. 877–889, 2011.

[30] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Anti-jamming games in multi-channel cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 4–15, jan 2012.

[31] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," *AAAI/IAAI*, vol. 1998, no. 746-752, p. 2, 1998.

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

[33] M. L. Littman, "Value-function reinforcement learning in markov games," *Journal of Cognitive Systems Research*, vol. 2, pp. 55–66, 2001.

[34] W. Liang, S. X. Ng, J. Feng, and L. Hanzo, "Pragmatic distributed algorithm for spectral access in cooperative cogni-

tive radio networks," *IEEE Transactions on Communications*, vol. 62, no. 4, pp. 1188–1200, 2014.

[35] W. Liang, S. X. Ng, and L. Hanzo, "Cooperative Overlay Spectrum Access in Cognitive Radio Networks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1924–1944, 2017.

[36] A. H. Bastami and P. Kazemi, "Cognitive multi-hop multi-branch relaying: Spectrum leasing and optimal power allocation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4075–4088, 2019.

[37] T. M. Chiwewe, C. F. Mbuya, and G. P. Hancke, "Using Cognitive radio for interference-resistant Industrial wireless sensor networks: An overview," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1466–1481, 2015.

[38] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications-a tutorial," *IEEE transactions on Communications*, vol. 30, no. 5, pp. 855–884, 1982.

[39] M. Grissa, B. Hamdaoui, and A. A. Yavuza, "Location privacy in cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1726–1760, 2017.

[40] K. Mourougayane and S. Srikanth, "Intelligent jamming threats to cognitive radio based strategic communication networks-a survey," in *2015 3rd International Conference on*

Signal Processing, Communication and Networking (ICSCN). IEEE, 2015, pp. 1–6.

[41] M. H. Ling, K.-L. A. Yau, J. Qadir, G. S. Poh, and Q. Ni, "Application of reinforcement learning for security enhancement in cognitive radio networks," *Applied Soft Computing*, vol. 37, pp. 809–829, 2015.

[42] S. Bhattacharjee, S. Sengupta, and M. Chatterjee, "Vulnerabilities in cognitive radio networks: A survey," *Computer Communications*, vol. 36, no. 13, pp. 1387–1398, 2013.

[43] D. Hlavacek and J. M. Chang, "A layered approach to cognitive radio network security: A survey," *Computer Networks*, vol. 75, pp. 414–436, 2014.

[44] H. Zhang, H. Xing, J. Cheng, A. Nallanathan, and C. M. Victor Leung, "Secure Resource Allocation for OFDMA Two-Way Relay Wireless Sensor Networks Without and with Cooperative Jamming," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1714–1725, 2016.

[45] Z. Lu, W. Wang, and C. Wang, "Modeling, Evaluation and Detection of Jamming Attacks in Time-Critical Wireless Applications," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1746–1759, aug 2014.

[46] O. Osanaiye, A. Alfa, and G. Hancke, "A Statistical Approach

to Detect Jamming Attacks in Wireless Sensor Networks," *Sensors*, vol. 18, no. 6, p. 1691, may 2018.

[47] F. Salahdine and N. Kaabouch, "Security threats, detection, and countermeasures for physical layer in cognitive radio networks: A survey," *Physical Communication*, vol. 39, p. 101001, 2020.

[48] M. A. Aref, S. K. Jayaweera, and E. Yepez, "Survey on cognitive anti-jamming communications," *IET Communications*, vol. 14, no. 18, pp. 3110–3127, 2020.

[49] G. Noubir, R. Rajaraman, B. Sheng, and B. Thapa, "On the robustness of ieee 802.11 rate adaptation algorithms against smart jamming," in *Proceedings of the fourth ACM conference on Wireless network security*, 2011, pp. 97–108.

[50] K. Pelechrinis, C. Koufogiannakis, and S. V. Krishnamurthy, "Gaming the jammer: Is frequency hopping effective?" in *2009 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*. IEEE, 2009, pp. 1–10.

[51] J. Zhang, K. Tan, J. Zhao, H. Wu, and Y. Zhang, "A practical SNR-guided rate adaptation," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, 2008, pp. 2083—-2091.

[52] T. D. Vo-Huu and G. Noubir, "Mitigating rate attacks through crypto-coded modulation," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2015, pp. 237–246.

[53] K. Pelechrinis, I. Broustis, S. V. Krishnamurthy, and C. Gkantsidis, "ARES: An anti-jamming reinforcement system for 802.11 networks," in *CoNEXT'09 - Proceedings of the 2009 ACM Conference on Emerging Networking Experiments and Technologies*, no. January, 2009, pp. 181–192.

[54] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*. IEEE, 2014, pp. 247–254.

[55] Ayaz Ahmad, Sadiq Ahmad, Mubashir Husain Rehmani and N. U. Hassan, "A Survey on Radio Resource Allocation in Cognitive Radio Sensor Networks," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 888–917, 2015.

[56] S. Alrabaee, A. Agarwal, D. Anand, and M. Khasawneh, "Game theory for security in cognitive radio networks," in *2012 International Conference on Advances in Mobile Network, Communication and Its Applications*. IEEE, 2012, pp. 60–63.

[57] Z. Song, X. Wang, Y. Liu, and Z. Zhang, "Joint spectrum resource allocation in noma-based cognitive radio network with swipt," *IEEE Access*, vol. 7, pp. 89 594–89 603, 2019.

[58] B. R. Deepak, P. S. Bharathi, and D. Kumar, "Radio frequency anti-jamming capability improvement for cognitive radio networks: An evolutionary game theoretical approach," in *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*. IEEE, mar 2017, pp. 1–6.

[59] J. He, C. Chen, S. Zhu, B. Yang, and X. Guan, "Antijamming game framework for secure state estimation in power systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2628–2637, 2018.

[60] I. K. Ahmed and A. O. Fapojuwo, "Stackelberg Equilibria of an Anti-Jamming Game in Cooperative Cognitive Radio Networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 1, pp. 121–134, 2018.

[61] M. Lichtman, J. D. Poston, S. Amuru, C. Shahriar, T. C. Clancy, R. M. Buehrer, and J. H. Reed, "A communications jamming taxonomy," *IEEE Security Privacy*, vol. 14, no. 1, pp. 47–54, 2016.

[62] Q. Zhou, Y. Li, and Y. Niu, "A countermeasure against ran-

dom pulse jamming in time domain based on reinforcement learning," *IEEE Access*, 2020.

[63] N. Van Huynh, D. T. Hoang, D. N. Nguyen, and E. Dutkiewicz, "Deepfake: Deep dueling-based deception strategy to defeat reactive jammers," *arXiv preprint arXiv:2005.07034*, 2020.

[64] D. Yang, G. Xue, J. Zhang, A. Richa, and X. Fang, "Coping with a smart jammer in wireless networks: A stackelberg game approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4038–4047, 2013.

[65] M. A. Aref and S. K. Jayaweera, "A cognitive anti-jamming and interference-avoidance stochastic game," in *2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*. IEEE, 2017, pp. 520–527.

[66] J. Baek, S. I. Han, and Y. Han, "Energy-efficient uav routing for wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 1741–1750, 2019.

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

[68] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan, "Stackelberg game approaches for anti-jamming defence in wireless

networks," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 120–128, 2018.

[69] H. Fang, L. Xu, and K.-K. R. Choo, "Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks," *Applied Mathematics and Computation*, vol. 296, pp. 153–167, mar 2017.

[70] A. H. Anwar, G. Atia, and M. Guirguis, "Adaptive topologies against jamming attacks in wireless networks: A game-theoretic approach," *Journal of Network and Computer Applications*, vol. 121, pp. 44–58, 2018.

[71] L. Yu, Q. Wu, Y. Xu, G. Ding, and L. Jia, "Power control games for multi-user anti-jamming communications," *Wireless Networks*, vol. 25, no. 5, pp. 2365–2374, 2019.

[72] L. Xiao, Y. Li, J. Liu, and Y. Zhao, "Power control with reinforcement learning in cooperative cognitive radio networks against jamming," *The Journal of Supercomputing*, vol. 71, no. 9, pp. 3237–3257, 2015.

[73] G. Dubosarskii, S. Primak, and X. Wang, "Multichannel power allocation game against jammer with changing strategy," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–5.

[74] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor,

"User-centric view of jamming games in cognitive radio networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2578–2590, 2015.

[75] Q. Chen, X. Hao, Z. Kong, and X. Yan, "Anti-sweep-jamming method based on the averaging of range side lobes for hybrid modulation proximity detectors," *IEEE Access*, vol. 8, pp. 33 479–33 488, 2020.

[76] F. Yao, L. Jia, Y. Sun, Y. Xu, S. Feng, and Y. Zhu, "A hierarchical learning approach to anti-jamming channel selection strategies," *Wireless Networks*, vol. 25, no. 1, pp. 201–213, 2019.

[77] J. Huang, G. Chang, and J. Huang, "Anti-jamming Rendezvous scheme for cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 648–661, 2017.

[78] S. Mneimneh, S. Bhunia, F. Vázquez-Abad, and S. Sengupta, "A game-theoretic and stochastic survivability mechanism against induced attacks in Cognitive Radio Networks," *Pervasive and Mobile Computing*, vol. 40, pp. 577–592, sep 2017.

[79] S. Misra, A. Mondal, P. Bhavathankar, and M.-S. Alouini, "M-jaw: Mobility-based jamming avoidance in wireless sensor net-

works," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5381–5390, 2020.

[80] Y. Gao, Y. Xiao, M. Wu, M. Xiao, and J. Shao, "Game theory-based anti-jamming strategies for frequency hopping wireless communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5314–5326, 2018.

[81] L. Jia, Y. Xu, Y. Sun, S. Feng, L. Yu, and A. Anpalagan, "A game-theoretic learning approach for anti-jamming dynamic spectrum access in dense wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1646–1656, 2018.

[82] L. Xiao, H. Dai, and P. Ning, "Jamming-resistant collaborative broadcast using uncoordinated frequency hopping," *IEEE transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 297–309, 2011.

[83] G. Han, L. Xiao, and H. Poor, "Two-dimensional anti-jamming communication based on deep reinforcement learning," in *International Conference on Acoustic Signal Speach Processing*, 2017, pp. 2087–2091.

[84] W. Xu, T. Wood, W. Trappe, and Y. Zhang, "Channel surfing and spatial retreats: defenses against wireless denial of ser-

vice," in *Proceedings of the 3rd ACM workshop on Wireless security*, 2004, pp. 80–89.

[85] G. K. Kurt and Ö. Cepheli, "Physical layer security of cognitive iot networks," in *Towards Cognitive IoT Networks*. Springer, 2020, pp. 101–123.

[86] K. Ding, X. Ren, D. E. Quevedo, S. Dey, and L. Shi, "Defensive deception against reactive jamming attacks in remote state estimation," *Automatica*, vol. 113, p. 108680, 2020.

[87] 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)*. IEEE, 2016, pp. 1–6.

[88] S. Guan, J. Wang, H. Yao, C. Jiang, Z. Han, and Y. Ren, "Colonel blotto games in network systems: Models, strategies, and applications," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 637–649, 2019.

[89] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using channel hopping to increase 802.11 resilience to jamming attacks," in *IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications*. IEEE, 2007, pp. 2526–2530.

[90] A. Garnaev, A. P. Petropulu, W. Trappe, and H. V. Poor,

"A jamming game with rival-type uncertainty," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5359–5372, 2020.

[91] Y. Yao, W. Zhou, B. Kou, and Y. Wang, "Dynamic spectrum access with physical layer security: A game-based jamming approach," *IEEE Access*, vol. 6, pp. 12 052–12 059, 2018.

[92] S. d'Oro, L. Galluccio, G. Morabito, S. Palazzo, L. Chen, and F. Martignon, "Defeating jamming with the power of silence: A game-theoretic analysis," *IEEE transactions on wireless communications*, vol. 14, no. 5, pp. 2337–2352, 2014.

[93] L. Xiao, J. Liu, Y. Li, N. B. Mandayam, and H. V. Poor, "Prospect theoretic analysis of anti-jamming communications in cognitive radio networks," in *2014 IEEE Global Communications Conference*. IEEE, 2014, pp. 746–751.

[94] E. Altman, K. Avrachenkov, and A. Garnaev, "A jamming game in wireless networks with transmission cost," in *International Conference on Network Control and Optimization*. Springer, 2007, pp. 1–12.

[95] L. Jia, Y. Xu, Y. Sun, S. Feng, L. Yu, and A. Anpalagan, "A multi-domain anti-jamming defense scheme in heterogeneous wireless networks," *IEEE Access*, vol. 6, pp. 40 177–40 188, 2018.

[96] H. Fang, L. Xu, and K.-K. R. Choo, "Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks," *Applied Mathematics and Computation*, vol. 296, pp. 153–167, 2017.

[97] 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 (ICNP)*, 2011.

[98] X. Wang, J. Wang, Y. Xu, J. Chen, L. Jia, X. Liu, and Y. Yang, "Dynamic spectrum anti-jamming communications: Challenges and opportunities," *IEEE Communications Magazine*, vol. 58, no. 2, pp. 79–85, 2020.

[99] M. Islam, S. Kandeepan, R. Evans *et al.*, "Multi-radio based rendezvous technique for heterogeneous cognitive radio sensor network," *Sensors*, vol. 21, no. 9, p. 2997, 2021.

[100] I. Elleuch, A. Pourranjbar, and G. Kaddoum, "A novel distributed multi-agent reinforcement learning algorithm against jamming attacks," *IEEE Communications Letters*, vol. 25, no. 10, pp. 3204–3208, 2021.

[101] C. Zhao, Q. Wang, X. Liu, C. Li, and L. Shi, "Reinforcement learning based a non-zero-sum game for secure transmission

against smart jamming," *Digital Signal Processing*, vol. 112, p. 103002, 2021.

[102] M. A. Shattal, A. Wisniewska, A. Al-Fuqaha, B. Khan, and K. Dombrowski, "Evolutionary game theory perspective on dynamic spectrum access etiquette," *IEEE Access*, vol. 6, pp. 13 142–13 157, 2017.

[103] Q. D. La, T. Q. Quek, J. Lee, S. Jin, and H. Zhu, "Deceptive Attack and Defense Game in Honeypot-Enabled Networks for the Internet of Things," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1025–1035, 2016.

[104] S. Nan, S. Brahma, C. A. Kamhoua, and N. O. Leslie, "Mitigation of jamming attacks via deception," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, pp. 1–6.

[105] D. T. Hoang, D. N. Nguyen, M. A. Alsheikh, S. Gong, E. Dutkiewicz, D. Niyato, and Z. Han, "' borrowing arrows with thatched boats': The art of defeating reactive jammers in iot networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 79–87, 2020.

[106] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Defeating reactive jammers with deep dueling-based decep-

tion mechanism," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.

[107] M. Zhu, A. H. Anwar, Z. Wan, J.-H. Cho, C. A. Kamhoua, and M. P. Singh, "A survey of defensive deception: Approaches using game theory and machine learning," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2460–2493, 2021.

[108] S. Bhunia, S. Sengupta, and F. Vázquez-Abad, "Performance analysis of CR-honeynet to prevent jamming attack through stochastic modeling," *Pervasive and Mobile Computing*, vol. 21, pp. 133–149, aug 2015.

[109] K. Firouzbakht, G. Noubir, and M. Salehi, "On the performance of adaptive packetized wireless communication links under jamming," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3481–3495, 2014.

[110] K. Bian, J.-M. Park, and R. Chen, "Control channel establishment in cognitive radio networks using channel hopping," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 689–703, 2011.

[111] A. Nasser, H. Al Haj Hassan, J. Abou Chaaya, A. Mansour, and K.-C. Yao, "Spectrum sensing for cognitive radio: Recent advances and future challenge," *Sensors*, vol. 21, no. 7, p. 2408, 2021.

[112] R. K. Mondal, B. Senadji, and D. Jayalath, "Dual-level sensing based multiple access protocol for cognitive radio networks," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017, pp. 1–5.

[113] S. Gamage, D. T. Ngo, and J. Y. Khan, "Statistical delay-qos driven resource allocation for multiuser cognitive radio networks," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[114] K. Ibrahim, S. X. Ng, I. M. Qureshi, A. N. Malik, and S. Muhaidat, "Anti-jamming game to combat intelligent jamming for cognitive radio networks," *IEEE Access*, vol. 9, pp. 137941–137956, 2021.

[115] T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

[116] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[117] L. Jia, F. Yao, Y. Sun, Y. Xu, S. Feng, and A. Anpalagan, "A hierarchical learning solution for anti-jamming stackelberg game with discrete power strategies," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 818–821, 2017.

APPENDIX A

Elements of Game Theory

The game theory began as an application of mathematics. It was employed during the Cold War to model and forecast Russian nuclear weapon motions. However, it later expanded into a discipline with a wide range of applications. Game theory is now widely applied in economics, sociology, politics, and engineering.

A.1 Preliminaries

Game theory simulates multi-agent interactions in which one agent's actions influence the outcomes of all other agents. In game theory, these agents are referred to as players. Set of $\mathcal{I}$ represents all players of the game. In N-player game, the set $\mathcal{I}_{-1}$ represent all opponents of the i^{th} player, where

$$\mathcal{I} = 1, 2, 3, \ldots, N \tag{A.1.1}$$

Each player has a collection of pure strategies that he can use to counteract his opponents. When you play against other players, it doesn't always imply you're hurting them. It's also possible to play cooperatively, in which each player help their opponents. The game's strategy space is the Cartesian product of all players' strategy spaces.

$$\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \times \cdots \times \mathcal{S}_N \tag{A.1.2}$$

Players choose their strategies at every turn of the game, and a strategy profile $s \in S$ is played. In game theory, if a player has a strategy of $s_i \in s$, the opponent's strategy is indicated as $s_{-i} = s/s_i$. As a result of the strategy profile s, each player is assigned a utility (payoff) $u_i(s_i, s_{-1} \in U_i$. All possible utilities U from all possible strategy profiles are collected in this set. U is another fundamental element of game theory.

$$\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{U}_3 \times \cdots \times \mathcal{U}_N \tag{A.1.3}$$

A.1.1 Mixed Strategy

A mixed strategy space $\sum_i$ is a probability distribution over the pure strategies S_i for every player i in the game. Therefore payoff of the player is the expected value of the pure strategies payoffs. And mixed strategy space of the game is the Cartesian product of individual

mixed strategy spaces similar to pure strategy space,

$$\sum = \sum_1 \times \sum_2 \times \sum_3 \times \cdots \times \sum_N \tag{A.1.4}$$

player i's payoff to mixed strategy profile σ is,

$$u_i(\sigma_i, \sigma_{-i}) = \sum_{s \in S} \left(\prod_{j=1}^{l} \sigma_j(s_j) \right) u_i(s_i, s_{-i}) \tag{A.1.5}$$

As a result, any pure strategy profile is a mixed strategy profile, with each player's probability weight accumulating on a single pure strategy. A game G is represented by

$$\mathcal{G} = < \mathcal{I}, \mathcal{S}, \mathcal{U}, \sum > \tag{A.1.6}$$

A.1.2 Pure Strategy

A thorough explanation of how a player will play a game is provided by a pure strategy. It determines the course of action a player will take in any given situation. The set of pure strategies available to a player is referred to as that player's strategy set.

A.2 Classification of Games

- **One shot game:** This is a game that can only be played once. The pay-off could be such that a game, such as mutually assured nuclear destruction, would be impossible to repeat. People are

often enticed to act opportunistically or selfishly in one-shot interactions.

- **Repeated games**: is the simplest kind of game in which each stage is repeated and usually these repetitions are indefinite, in infinite time horizon. Let N denote the number of secondary users in a game and a^k represents the set of actions taken by those N SUs in kth stage of the game. The action in each stage k is to maximize utility by taking into account the history of the actions taken in previous stages. The expected strategy is discounted by the factor $0 < \delta < 1$, which means that the payoff of the current stage is worth greater than trailing payoffs. The simple example, in this case, may be taken as a jamming scenario. Where N secondary users are the players of the game. Among these N players J are jammers and K are legitimate secondary users such that $N = K + J$. Their respective actions are to choose the C available channel by avoiding the jammer and to choose C_j channels to jam the channels to reduce the payoff of the legitimate users.

- **Stochastic game:** A stochastic game is a recurring game with probabilistic transitions performed by one or more players, first introduced by Lloyd Shapley in the early 1950s. The game is divided into several stages. The game is in some state at the start of each stage. The players choose actions, and each receives

a reward based on the current condition and the acts taken. The game then shifts to a new random state, the distribution of which is determined by the previous state and the players' activities. Play continues for a finite or infinite stages after the method is repeated in the new state.

A.3 Equilibrium Concepts

A solution concept in game theory is a formal rule for predicting how a game will be played. These predictions are known as 'solutions,' and they describe which strategies players will use and, as a result, the game's outcome. Most famously Nash equilibrium, Equilibrium notions are the most widely employed solution concepts.

A.3.1 Nash Equilibrium

The Nash equilibrium is a concept in game theory that states that the best outcome of a game is one in which no player has an incentive to deviate from their chosen strategy after considering the strategy of an opponent. According to Nash equilibrium, the ideal conclusion of a game occurs where there is no incentive to depart from the beginning strategy, a notion in game theory.

A.3.2 Stackelberg Equilibrium

The Stackelberg equilibrium model is a strategic game in which the leader player moves first, followed by the follower player. It was named after the German economist Heinrich Freiherr von Stackelberg, who presented the concept in his 1934 book Market Structure and Equilibrium (Marktform und Gleichgewicht). In terms of game theory, the game's players are leaders and followers who compete on quantity.