MULTI-HOP COMMUNICATION IN COGNITIVE RADIO NETWORKS: AN EXPERIMENTAL EVALUATION

by

SYED AQEEL RAZA

Submitted in total fulfilment of the requirements for the

Ph.D. (Computing)

at Sunway University

August, 2016

DEDICATED TO MY PARENTS AND LATE ABBI (UNCLE)

ACKNOWLEDGEMENT

In the Name of Allah, the Most Beneficent, the Most Merciful; and I am thankful to Him for His countless blessings on me, through which I am able to conduct this research and complete this thesis. I would like to express gratitude and special acknowledgement to my mentor and supervisor Dr. Yau Kok Lim for his endless support, expert guidance, and patience in dealing with me during the entire duration of my study and research.

Special thanks to Sunway University that offered me the full scholarship and provides me an opportunity to work as teacher assistant in this great alma meter; along with this I am thankful to all its staff, members and colleagues for their continuous support and favor. I would also like to thank Dr. Hafizal Mohammad and Dr. Nordin Ramli for their help and cooperation during my stay in MIMOS as a research assistant.

Finally, I would like to thanks to my parents, family and friends for their unconditional support, love and prayers during the tenure of my education; I would not have been able to achieve this milestone without their encouragement that give me a hope and comfort.

ABSTRACT

Cognitive radio (CR) is the next-generation wireless communication system that has been proposed to address spectrum scarcity in the traditional static spectrum assignment policy. The static spectrum assignment policy allocates frequency bands to licensed users or primary users (PUs) for their exclusive usage, and so unlicensed users or secondary users (SUs) are forbidden from accessing the licensed channels. CR uses dynamic spectrum access to solve this problem, which have two main approaches to access channels. Firstly, SUs access the channel in an opportunistic manner, while the PUs are oblivious to the presence of SUs. Secondly, SUs negotiate with PUs for channel access in a collaborative manner in order to achieve mutual benefit, such as Quality of Service (QoS) enhancement for both PUs and SUs.

To date, research has been primarily focused on simulation-based investigation. There were only a perfunctory effort to investigate the network layer of CR networks on a real testbed environment. This research work is a pioneering effort to examine opportunistic and collaborative channel access approaches at the network layer using real testbed implementation. There are three major contributions in this thesis. Firstly, a channel selection scheme is implemented in multi-hop CR network using reinforcement learning (RL) on a USRP/ GNU radio platform. Secondly, route selection schemes are implemented in a multi-hop CR network using RL and SL with the objective of improving QoS performance. Thirdly, addresses the challenges of network-layer implementation using USRP/ GNU radio platform. Analyzes the outcomes and results of the proposed schemes implemented on a USRP/ GNU radio platform.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

1. Syed AR, and Yau KLA. Spectrum Leasing in Cognitive Radio Networks. International Journal of Distributed Sensor Networks, vol. 2014, DOI: 10.1155/2014/329235, pp. 1-22, 2014. [ISI IF 2013 (0.923), Scopus].

2. Syed AR, Yau KLA, Mohamad H, Ramli N, and Hashim W. Channel Selection in Multi-hop Cognitive Radio Network using Reinforcement Learning: An Experimental Study. IET International Conference on Frontiers of Communications, Networks and Applications (ICFCNA'14) IEEE, Kuala Lumpur, Malaysia, November 2014.

3. Syed AR, and Yau KLA. Relay Node Selection for Spectrum Leasing in Cognitive Radio Networks. IEEE International Conference on Control System, Computing and Engineering, Penang, Malaysia, November 2013.

4. Syed AR, Yau KLA, Qadir J, Mohamad H, Ramli N, and Keoh SL. Route Selection for Multi-hop Cognitive Radio Networks using Reinforcement Learning: An Experimental Study. IEEE Access, vol. 4, DOI: 10.1109/ACCESS.2016.2613122, pp. 6304-6324, 2016. [ISI IF 2015 (1.249)]

5. Syed AR, and Yau KLA. On Cognitive Radio-based Wireless Body Area Networks for Medical Applications. IEEE Symposium Series on Computational Intelligence (SSCI) IEEE, Singapore, April 2013.

AWARDS AND MEDALS

1. "Multi-hop Cognitive Radio Networks: The Next Generation Wireless Communication Systems," 27th International Invention, Innovation and Technology Exhibition, Silver Medal, May 2016.

2. "A Security Enhanced Routing Mechanism with Anti-jamming Technique for Multi-hop Networks using Cognitive Radio," 26th International Invention, Innovation and Technology Exhibition, Silver Medal, May 2015.

3. "Cognitive Radio Technology with Networking with Networking and Security Enhancements applied to Wireless Surveillance Systems," 25th International Invention, Innovation and Technology Exhibition, Gold Medal, May 2014.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CRN	Cognitive radio network
CR	Cognitive radio
PUs	Primary users
SUs	Secondary users
QoS	Quality of Service
RL	Reinforcement learning
USRP	Universal software radio peripheral
DSA	Dynamic spectrum access
ISM	Industrial, scientific and medical
SL	Spectrum leasing
BS	Base Station
SNR	Signal-to-noise ratio
TDMA	Time-Division Multiple Access
SU BS	Secondary user base station
PU BS	Primary user base station
FCC	Federal communications commission
BER	Bit error rate
RTS	Request to send
RTC	Request to cooperate
CTC	Clear to coordinate
CTS	Clear to send

OFDM	Orthogonal frequency division multiplexing
MDP	Markov decision processes

1.0 INTRODUCTION

Cognitive radio (CR) is the next-generation wireless communication system that addresses the problem of spectrum scarcity [1]. While the CR concept has been getting popular recently, the industry has not yet adopted it as a platform of wireless communication on a commercial basis. With the advent of testbed platforms, some initial implementation work has been carried out on a small scale, which may turn into a first step towards the complete implementation of CR network in near future. This thesis presents the implementation of CR network on a real testbed platform in an environment of multi-hop communication, and specifically addresses the network layer implementation, which is still in its infancy and developmental stage. The key aim of this thesis is to improve the Quality of Service (QoS) performance of CR network.

The rest of this chapter provides a general overview of this research. Specifically, Section 1.1 presents the motivation. Section 1.2 presents the research objectives. Section 1.3 presents the problem statements. Section 1.4 presents the research questions. Section 1.5 presents the research contributions. Section 1.6 presents the thesis outline.

1.1 Motivation

The fast and steady growth of new wireless technologies, such as IEEE 802.11 [2], IEEE 802.16 [3], and Long Term Evaluation [4] and advanced-LTE [5], has caused increasing stress on the limited spectrum availability as a result of the traditional static spectrum assignment policies adopted by the telecom regulatory authorities worldwide. For instance, Malaysian Communications and Multimedia Commission is a telecom regulatory body in Malaysia. The static spectrum assignment allocates frequency bands, and hence exclusive channel access, to licensed users; and so unlicensed users are forbidden from accessing the licensed channels. Previous studies revealed that in the exclusive use of spectrum channels by PUs, the spectrum utilization is in the range of 15% to 85% [6]. This has led to severe spectrum scarcity, which creates bottlenecks to support a large amount of packet transmissions among unlicensed users. Nevertheless, the main reason of spectrum scarcity is the static spectrum assignment, rather than the utilization of spectrum at full capacity [6].

Cognitive radio has been proposed to address this problem through Dynamic Spectrum Access (DSA). In CR networks, there are two types of users, namely licensed users (or Primary Users, PUs) and unlicensed users (Secondary Users, SUs). Generally speaking, DSA can be categorized into three types: open sharing, hierarchical access and dynamic access [7]. The open sharing approach enables all nodes to access the channels in an equal manner; and this approach has been adopted in unlicensed bands such as Industrial, Scientific and Medical (ISM) bands [8]. The hierarchical access segregates the nodes in a hierarchical manner in which the PUs are given higher priority to access their respective channels compared to SUs, and so the SUs must relinquish their access whenever PUs' activities reappear. Hence, the hierarchical access enables SUs to opportunistically access the channels while PUs is in the idle state. This approach has been adopted in the traditional opportunistic access schemes in CR networks [1]. The dynamic access enables entities in the network to negotiate for channel access in a collaborative or partnership manner in order to achieve mutual benefit, such as Quality of Service (QoS) enhancement, at most entities. This approach has been adopted in spectrum leasing (SL) [9], in which SUs negotiate with PUs for the acquisition of underutilized channels, in which PUs allocate the channel access time to SUs. This is a win-win solution for both PUs and SUs. For instance, the main advantage for PUs is that, the PUs may either enhance its QoS performance (e.g. throughput) by relaying the PUs' packets towards its destination through SUs, or to increase its monetary gain by leasing its spectrum to SUs. Whereas, the main advantage for SUs is that, the SUs enhance its QoS performance by obtaining guaranteed channel access (rather than opportunistic access) from PUs.

Generally speaking, the research in wireless networks is evaluated using three models; that are mathematical model, simulation model, and experimental model, respectively. In mathematical model, the investigation is relied on mathematical formalisms to probe the research problems [10]. In simulation model, the investigation is relied on a computer based simulation tools (e.g, OMNET++, Qualnet) [11]. Lastly, in experimental model, the investigation is relied on a real testbed implementation (e.g., USRP/ GNU radio) [12]. Tremendous research work has been done to investigate the environment of cognitive radio (CR) network as well as spectrum leasing (SL) using mathematical model and simulation model at all the layers of OSI network model, including the network layer [13]. For instance, in recent years some of the research related to the mathematical model of CR network is presented in [14-19], while some research related to the simulation model of CR network is presented in [20-25]. However, perfunctory efforts have been made to investigate the environments of CR network and SL on an experimental model at network layer, and has been presented in [26-28]. Cognitive Radio and Spectrum leasing are proposed in this thesis to achieve some of the main objectives, specifically, implementing the proof of concept prototype on a real testbed environment, and subsequently to enhance the QoS of SUs. The incorporation of intelligence using the reinforcement learning (RL) technique [29-31] into the proposed schemes on a real testbed implementation is one of the distinguish feature of this thesis. By implementing the proof of concept prototype on a real testbed environment, the CR and SL are envisioned to be applied in a wide range of new applications [32], such as to enhance the Quality of Service (QoS) performance in wireless body area networks [33].

1.2 Research objectives

The research objectives are as follows:

1. To identify the research problems by exploring the existing spectrum leasing (SL) and real testbed implementation schemes in cognitive radio network (CRN).
2. To propose and design:
   i) A simple SL approach in which PUs select SUs as relay nodes, on a real testbed environment, as well as using a simulation tool.
   ii) A channel selection mechanism is implemented in a multi-hop CRN.
   iii) Simple and novel CR and SL schemes as a proof of concept on a real testbed environment. A route selection mechanism is implemented on a multi-hop CRN using different network topologies.
3. To investigate integration of reinforcement learning (RL) as an artificial intelligence technique. The RL approach is known to sense the operating environment, as well as to learn and relearn its knowledge about the operating environment.

Generally speaking, the objective is to investigate the above mentioned schemes while incorporating the RL as an artificial intelligent technique in order to enhance the QoS performance on a real testbed environment.

1.3 Problem statement

Generally speaking, many challenges are associated with CR networks. For instance, SUs need to select channels for their respective packet transmission with their neighboring nodes in order to enhance the QoS performance. This challenge is becoming more complicated in multi-hop communication among SUs nodes having different routes from SU source node to SU destination node. Secondly, the selection of channel/ route in a multi-hop communication is difficult due to the time-varying activities of PUs. Thirdly, the implementation of the aforementioned challenges on a real testbed is still at very infancy stage. Therefore on a real testbed environment, multi-hop communication requires more intelligence for the selection of channel/ route due to the activities of PUs' that may appear and disappear from time to time. The main goal of this thesis is to enhance the QoS performance of SU network; while reducing the number of channel switching/ route breakage in a multi-hop communication on a real testbed environment.

1.4 Research questions

This thesis investigates the following four research questions:

1. What are the recent advances in spectrum leasing for CR networks and the real testbed implementation for the deployment of multi-hop communication in CR network?

2. How to design the system model that covers PUs' activities and SUs architecture, which is suitable and scalable for real testbed implementation?
3. How to apply reinforcement learning (RL) approach on a real testbed implementation?
4. How to improve the QoS performance of SUs in the proposed schemes using the centralized/ distributed models?

1.5 Research contributions

The contributions are as follows:

1. Implements a channel selection scheme in multi-hop CR network using RL on a USRP/ GNU radio platform.
2. Implements the route selection from a SU source node to a SU destination node in a multi-hop network using RL and SL with the objective of improving QoS performance taking into consideration the challenges of the underlying USRP/ GNU radio platform.
3. Addresses the challenges of network-layer implementation using USRP/ GNU radio platform. Analyzes the outcomes and results of the proposed schemes implemented on a USRP/ GNU radio platform.

1.6 Thesis outline

The rest of this thesis is divided into the following chapters.

Chapter 2 reviews three main research areas, namely cognitive radio, spectrum leasing and the limited real testbed implementation of CRN. It provides overviews on CR and spectrum leasing, as well as discusses various aspects including the advantages, characteristics, challenges, performance enhancement achievements and open issues. Finally, it presented the limited state-of-the-art work related to the implementation of CRN on a real testbed environment.

Chapter 3 discusses the preliminary evaluation and results using a simulation and testbed environments.

Chapter 4 presents the real testbed implementation for the deployment of channel selection schemes using multi-hop CR network with the help of USRP/ GNU radio.

Chapter 5 presents the real testbed implementation for the deployment of route selection schemes using multi-hop CR network with the help of USRP/ GNU radio.

Chapter 6 presents summary and future work.

2.0 LITERATURE REVIEW

This chapter presents the state-of-the-art literature review of CR, SL and investigate the limited work on a real testbed implementation, specifically at the network layer. The state-of-the-art CR and SL schemes presented in this chapter are either simulated using the simulation model or solved using the mathematical model, which highlights the need to implement the concept of CR and SL on a real testbed environment. There are five sections in this chapter. The first two sections (i.e., section 2.1 and section 2.2) provide overviews on CR and spectrum leasing, as well as discuss various aspects including the advantages, characteristics, challenges, and performance enhancement achievements. Section 2.3 presents an open issues in SL for CRN. Section 2.4 presents the limited state-of-the-art work implemented on the real testbed environment, specifically at the network layer. Finally, section 2.5 presents the chapter summary.

2.1 Cognitive radio

This section presents an introduction of CR, as well as its characteristics, advantages and disadvantages.

2.1.1 Introduction

The revolution of ubiquitous and pervasive computing has significantly increased the demand for spectrum resulting in spectrum scarcity. CR network is the next generation wireless communication system, has emerged as a promising technique to address spectrum scarcity through dynamic spectrum access. Traditionally, the PUs are the licensed users, and they have exclusive right to use their channels. Whereas, the *CR* network enables *opportunistic channel access*, in which SUs access the underutilized channels (or white space) opportunistically whenever PUs are in an idle state. The SUs can access licensed and unlicensed bands in order to improve spectrum efficiency [6]. Furthermore, SUs are oblivious to PUs, and therefore the acquisitions of white spaces are not guaranteed. In licensed bands, PUs have higher priority to access the channels than SUs; while in unlicensed bands, there is lack of the concept or identities of PUs and SUs such that every user has the same priority to access the channels [1]. Figure 2.1 shows the opportunistic channel access by SUs at different time when PUs are in their idle state.

2.1.2 Characteristics and advantages of CR

Cognitive Radio has unique attributes of '*learn*', '*sense*', and '*adapt*'. In CR, SUs are equipped with the capability to *observe* and *learn* from the operating environment, so SU

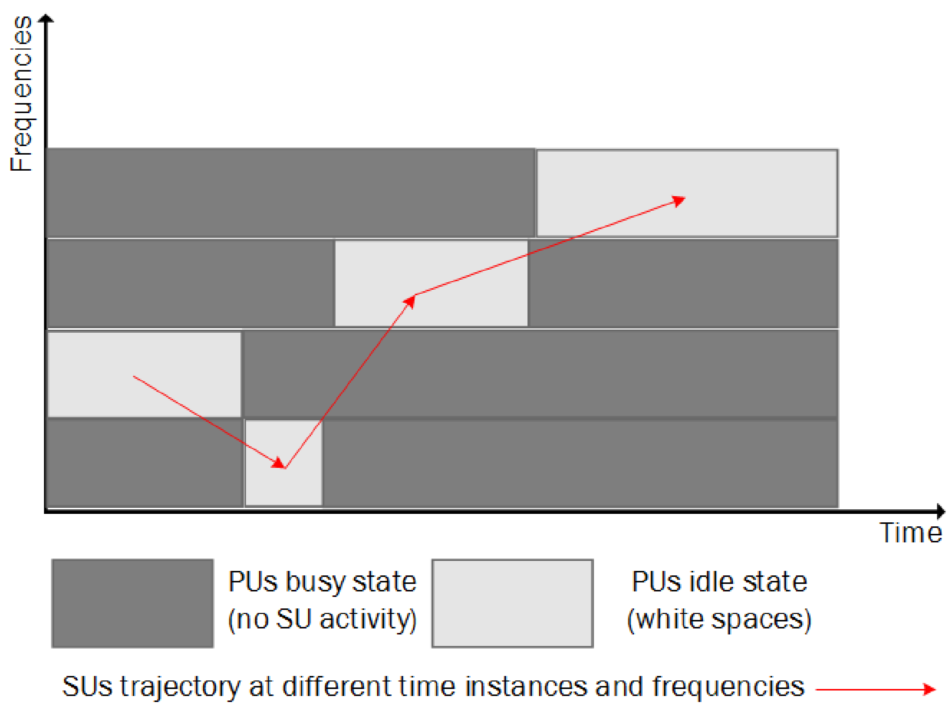


Figure 2-1: Opportunistic channel access by SUs

transceivers can adapt their respective wireless transmission parameters and make decisions to optimize the network performance [14]. These aforementioned characteristics contribute to the advantages of CR. Generally speaking, CR provides two main advantages. Firstly, cognition capabilities enable SUs to acquire appropriate information about transmission parameters (e.g., transmission power) through learning mechanism, and to capture the unused portion of spectrum through sensing the surrounding operating environment. Secondly, re-configurability enables SUs to dynamically adjust their transmission parameters (e.g., transmission power) respectively through adaptation mechanism based on the surrounding operating environment [14].

2.1.3 Shortcomings of CR

There are two shortcomings associated with the traditional CR networks. Spectrum leasing (see section 2.2) has been proposed to overcome these shortcomings.

2.1.3.1 Channel uncertainty due to reappearance of PUs

SUs acquire the channel in an opportunistic manner whenever the PUs are in an idle state or the PUs are physically absence. During the absence of PUs, SUs can use the underutilized channels to improve their network performance. However, whenever the PUs reappear in the channels, SUs must evacuate the channels on an immediate basis. This evacuation of channel may severely deteriorate the SUs' network performance due the uncertainty in the reappearance of PUs' activities [34].

2.1.3.2 Lack of PUs' reward

In CR networks, only the SUs gained rewards by acquiring the channels for their use, and this improves the SUs' network performance. Hence, since the PUs do not receive any rewards, there has been lack of encouragement for the PUs to share their channels with SUs [35].

2.2 Spectrum leasing

Since a SU must vacate its channel whenever a PU reappears in the channel, this may affect the SUs' network performance [34]. In view of this shortcoming, spectrum leasing has been proposed, in which the PUs and SUs negotiate with each other in a manner that allows SUs to acquire white spaces for a guaranteed period of time, which is a win-win solution for both PUs and SUs [9]. Through spectrum leasing, the PUs enhance their network performance and maximize their monetary gain; while the SUs enhance their network performance [17]. While there are numerous research efforts investigating CR, the research into spectrum leasing remains at its infancy. This section, present a comprehensive review on spectrum leasing schemes in CR networks by highlighting some pioneering approaches in this area; and it covers the discussion on the gains, functionalities, characteristics and challenges of each scheme in CR networks. Additionally, this section discusses performance enhancement achieved by the spectrum leasing schemes, as well as various open issues in order to spark new interests in this research area.

2.2.1 Introduction

Spectrum leasing is a dynamic spectrum access technique in which PUs and SUs form a partnership for mutual benefits. In spectrum leasing, the SUs negotiate with PUs and acquire their white spaces [9], while the PUs lease their channels and receive rewards in the form of monetary gain or network performance enhancement through packet forwarding by SUs [36]. Hence, PUs are fully aware of the presence of SUs. Figure 2.2 presents a taxonomy of spectrum leasing, which covers its advantages, functionalities, characteristics and challenges. Further descriptions about the taxonomy are found in the later subsections. Generally speaking, with the use of spectrum leasing, PUs and SUs receive the following advantages represented by A1 and A2 (see Figure 2.2), respectively:

A1 PU's gain

A1.1 *Monetary gain*. PUs may lease its licensed channels during idle periods for financial reward or revenue. For instance in [9], Jayaweera et al. (2010) propose a PU's utility function based on its monetary gain (e.g. the price set by PUs of white spaces).

A1.2 *Network performance enhancement*. The PU links may deteriorate due to shadowing and interference. Through spectrum leasing, one or more SUs form an alternative route and relay PUs' traffic, and this enhances the PUs' network performance, such as successful transmission rate, throughput, end-to-end delay and energy efficiency [37].

A2 SU's gain

A2.1 *Dedicated channel access*. The SUs access white spaces allocated by PUs. Subsequently, this enhances the SUs' throughput performance. Since spectrum leasing enhances the throughput performance of PUs (A1.2), it reduces the transmission time of PUs, therefore leaving more white spaces and transmission opportunities to SUs for dedicated access [38].

The advantages motivate PUs and SUs to participate in spectrum leasing. For instance, in [39], spectrum leasing maximizes a weighted sum of PUs' and SUs' throughput performance.

This section provides an extensive survey on existing spectrum leasing schemes in CRNs. The purposes are to establish a foundation and to spark new interests in this

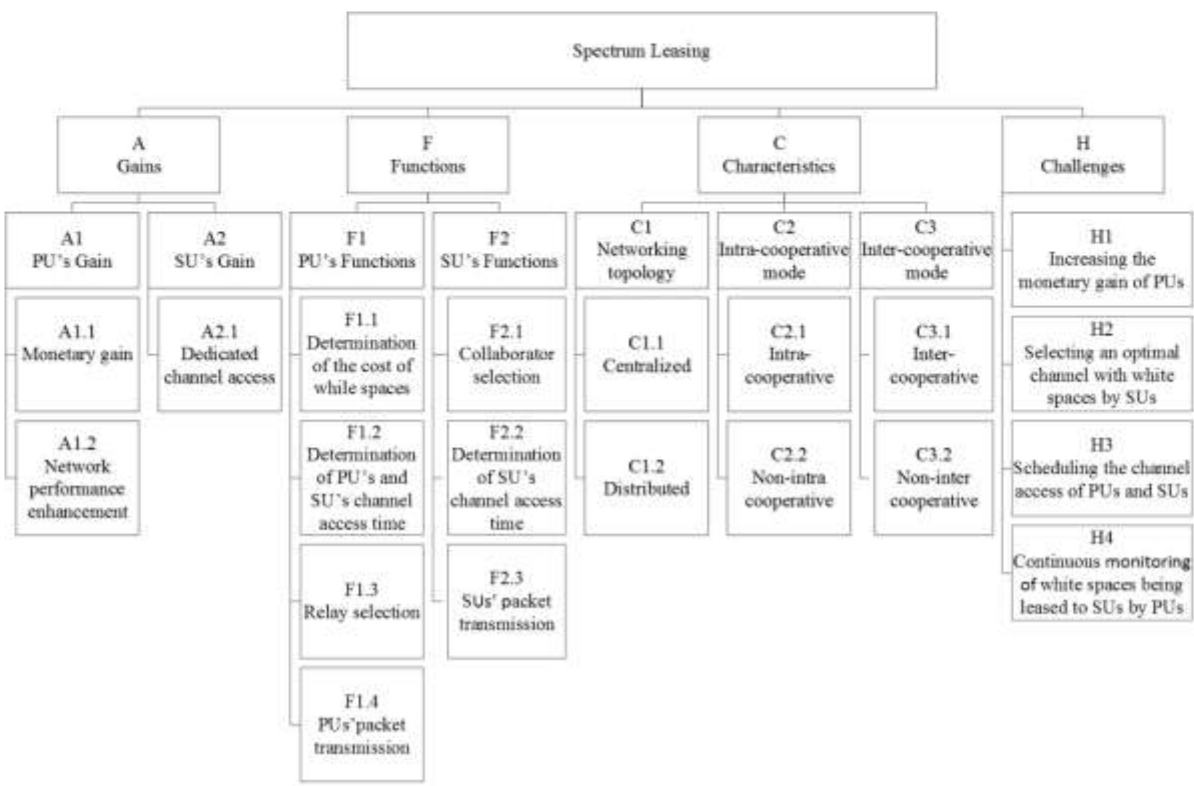


Figure 2-2: Taxonomy of spectrum leasing in CRNs

research area covering new kinds of CR networks such as CR sensor networks [40]. The contributions are as follows. The below subsections, present the functionalities, characteristics and challenges, followed by various spectrum leasing schemes in CRNs, and performance enhancement achieved by spectrum leasing schemes.

2.2.2 Functionalities of spectrum leasing in cognitive radio networks

This subsection discusses the functionalities of PUs and SUs for spectrum leasing in CRNs. Generally speaking, spectrum leasing is comprised of the following functionalities:

F1 PU's function

F1.1 *Determination of the cost of white spaces*. PUs determine the cost (e.g. monetary price) of white spaces to be imposed on SUs.

F1.2 *Determination of PUs' and SUs' channel access time.* PUs are the rightful owners of the licensed spectrum, and so the PU Base Station (BS) may determine suitable channel access time for transmission opportunities for both PUs and SUs. For instance, in centralized networks, the PU hosts send their respective information (e.g. idle time) to PU BS. Subsequently, the PU BS allocates transmission opportunities for PU and SU networks. In other words, the PUs determine the amount of white spaces to be leased to SUs. The objective is to maximize the network performance (e.g. throughput) of PUs and SUs [42, 43].

F1.3 *Relay selection*. PUs select the SUs that provide the highest gain (e.g. PU-SU links with the best-known Signal-to-Noise Ratio (SNR)) as relays in order to maximize throughput performance.

F1.4 *PUs' Packet transmission*. PUs transmit their own packets to destination in order to enhance their network performance.

F2 SU's function

F2.1 *Collaborator selection.* SUs select the suitable PUs to collaborate with. This covers the evaluation of the gain (e.g. the amount of white spaces with sufficient SNR) and cost (resources required to relay PUs' traffics, such as energy consumption).

F2.2 *Determination of SU's channel access time.* SUs determine the amount of white spaces, which increases with channel access time, to request from PUs based on the cost imposed by the PUs. For instance, in a Time-Division Multiple Access (TDMA) system, SUs must determine the optimal time duration in which they must involve as relay to transmit PU packets and to transmit their own packets [37].

F2.3 *SUs' Packet transmission*. SUs transmit packets, and this involves two phases. Firstly, the SUs relay PU packets. To ensure continuous collaboration with PUs, the SUs must achieve a certain level of network performance enhancement while relaying the PUs' packets. Secondly, the SUs transmit their own packets. Spatial reuse is possible, and so the SUs must minimize interference among themselves [44]. For instance, in centralized networks, SU BS and hosts may serve as relays to transmit PU packets, and subsequently the SU BS allocates the white spaces offered by PUs to its SU hosts fairly [42, 45].

Spectrum leasing involves several steps and message handshaking, and Figure 2.3 describe a general procedure. Consider two centralized PU and SU networks, which are collocated in the same area. Several PU hosts (or SU hosts) are associated with a PU BS (or SU BS). The procedure is as follows:

Step 1. The PU hosts send information on their respective idle periods (or white spaces) to PU BS.

Step 2. The PU BS determines the cost (F1.1) and duration (F1.2) of white spaces. There are J PU hosts to be leased to SUs.

Step 3. The PU BS sends the cooperation information (e.g. the cost and duration, as well as SNR of the white spaces) to SU BS.

Step 4. The SU BS broadcasts the cooperation information to its SU hosts.

Step 5. The SU hosts determine the optimum transmission and relaying strategies (i.e. F2.2 and F2.3) using the cooperation information. If auction mechanism is applied, the SU hosts may determine bid values.

Step 6. The SU hosts send their respective decisions (e.g. strategies and bid values) to SU BS.

Step 7. The SU BS decides to accept the lease or not, and select the suitable PUs to collaborate with (F2.1).

Step 8. The SU BS sends its decisions to PU BS.

Step 9. The PU BS decides to lease or not, and select the suitable SUs as relays (F1.3).

Step 10. Finally, based on the lease, the PU BS transmits its packets (F1.4) directly through a single hop, or indirectly through SU relay nodes, to the PU BS's destination node. The SU BS may divide the white spaces and assign the access time of each white space to each SU hosts (F2.2). The SUs transmit packets accordingly (F2.3).

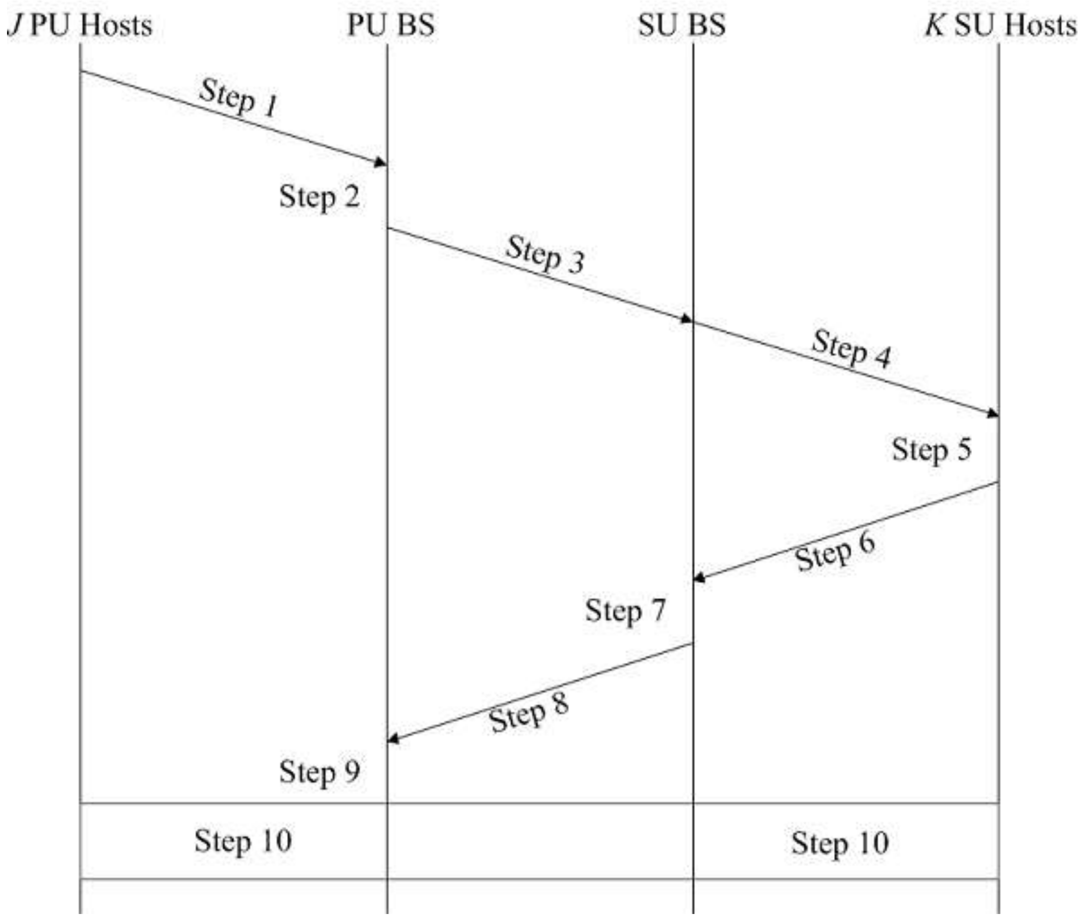


Figure 2-3: A general spectrum leasing procedure

2.2.3 Characteristics of spectrum leasing in cognitive radio networks

This subsection discusses the characteristics of spectrum leasing in CRNs. There are three characteristics as follows:

C1 *Network topology: Centralized* (C1.1) and *Distributed* (C1.2). In centralized networks (C1.1), a central entity which is usually referred as Base Station (BS) is responsible for communications between PU and SU networks [39]. Whereas, in distributed networks (C1.2), BS does not exists, and PUs and SUs share their information through a common control channel [46]. For instance, in [39], a centralized network (C1.1) topology is used, in which PUs are leaders and responsible to select the most appropriate SU for cooperative communication; and hence the SUs are followers.

C2 *Intra-cooperative mode: Intra-cooperative* (C2.1) and *Non-intracooperative* (C2.2). The PUs may cooperate among themselves through an intra-cooperative approach in order to achieve the advantages (A1.1)–(A1.2) and (A2.1). Likewise, the SUs may adopt the same approach. In Figure 2.4, the intra-cooperative (C2.1) mode is shown in (a) and (c), from the SU's perspective the SUs may cooperate among themselves and jointly improve network-wide performance such as throughput performance, as well as to reduce the monetary and non-monetary spectrum leasing costs imposed by PUs. In other words, a group of SUs may lease a channel, and subsequently share the channel among themselves in order to reduce spectrum leasing costs. In Figure 2.4, the non-intracooperative (C2.2) mode is shown in (b) and (d), from the PU's perspective each PU may compete with each other to lease their respective white spaces; and hence, each PU may set a competitive price based on the demand of channel access from SUs. From the SU's perspective the SUs may also compete with each other to acquire the white spaces through auction-based mechanisms [47]. For instance, in [39], each SU optimizes its power allocation in the transmission of PU packets in order to fulfill the packet transmission requirements of PUs. This helps each SU to remain competitive in order to obtain white spaces in the upcoming auctions; and this has been shown to improve SU throughput performance.

C3 *Inter-cooperative mode: Inter-cooperative (C3.1) and Non-intercooperative (C3.2).* A PUs and SUs may cooperate with each other in order to achieve the advantages (A1.1)–(A1.2) and (A2.1). In Figure 2.4, the inter-cooperative (C3.1) mode is shown in (c) and (d), the PUs and SUs cooperate with each other, and so this improves the overall network-wide performance such as throughput performance. In Figure 2.4, the non-intercooperative (C3.2) mode is shown in (a) and (b), the PUs and SUs are referred as selfish users, and they do not cooperate with each other. For instance, in [48], the PUs attempt to maximize their profit or reward out of the white spaces; while the SUs attempt to reduce their cost.

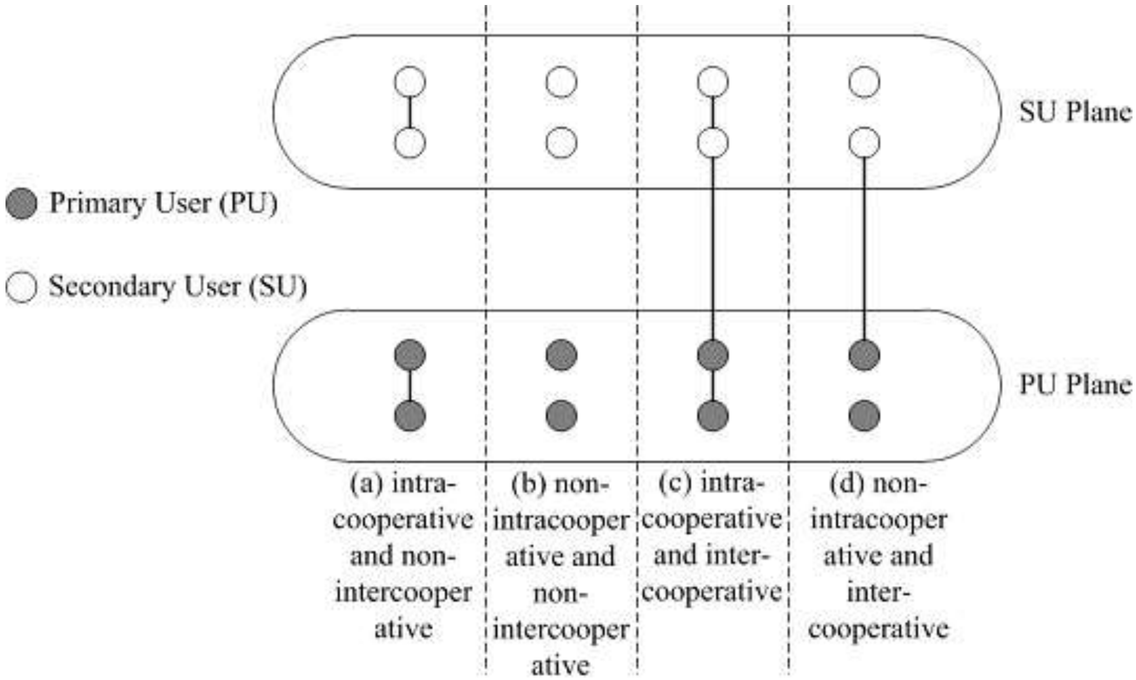


Figure 2-4: Mode of cooperation between PU and SU network

2.2.4 Challenges of spectrum leasing in cognitive radio networks

This subsection discusses the challenges associated with spectrum leasing in CRNs. There are four challenges as follows:

H1 *Increasing the monetary gain of PUs*. PUs aim to increase their monetary gain through spectrum leasing. This encourages the PUs to participate in spectrum leasing by increasing the amount of white spaces available to SUs. Subsequently, this increases PUs' and SUs' throughput performance [42]. The PUs may cooperate or compete with each other to lease their white spaces. As an example, in [42], PUs cooperate with each other, and linear programming is applied to set the optimal price of the white spaces in order to increase their monetary gain. As another example, in [48], PUs compete with each other, and game theory is applied to set the optimal price of the white spaces in order to increase their monetary gain.

H2 *Selecting an optimal channel with white spaces by SUs.* SUs aim to access the licensed channel or white spaces in order to increase their network performance (e.g. throughput). So, this encourages the SUs to participate in spectrum leasing, and subsequently increases PUs' and SUs' network performance [49]. However, the access to white spaces by SUs requires monetary cost, and so there is a need to find an optimal channel that provides the best possible network performance while incurring the least possible cost. For instance in [39], Cao et al. (2012) propose a spectrum sharing policy in which white spaces being leased to SUs, in order to increase the network capacity of SU network.

H3 *Scheduling the channel access of PUs and SUs.* The PUs schedule the time for the transmissions of PUs' and SUs' packets in order to enhance their respective QoS performance (e.g. throughput). The time allocation for SUs' links must be sufficiently higher compared to that of PUs' links in order to reap the benefits of spectrum leasing [38]. Otherwise, the queue size at SU relay nodes may grow, and eventually insufficient to accommodate new packets from both PUs and SUs leading to packet loss. However, the white spaces being leased to SUs may not be sufficient to cater for PUs' and SUs' packets. For instance, in [50] Huang et al. (2011) propose a coalition game to allocate a suitable fraction of channel access time among PUs and SUs, so that SUs transmit PUs' packets as well as their own packets.

H4 *Continuous monitoring of white spaces being leased to SUs by PUs.* Upon negotiation, the PUs and SUs may need to monitor the white spaces (e.g. amount and channel quality) and the Quality of Service (QoS) of packet transmission in order to make sure that each party follows suit. However, the continuous monitoring of SUs requires more intelligence to be incorporated into the PU network. For instance, in [47], PUs additionally acts as an online auctioneer to monitor the SUs activities. Likewise, in [51], PUs need to ensure that the interference caused by SUs is less than the acceptable interference level. Furthermore, SUs also need to monitor the SUs' signal level in order to reduce interference with PUs [52].

2.2.5 Spectrum leasing schemes in cognitive radio networks

This subsection presents existing work on spectrum leasing schemes in CRNs. The schemes are categorized with respect to the challenges (see subsection 2.2.4) and on the basis of adopted approaches (e.g. game theoretic approaches and non-game theoretic approaches) to address the challenges. The game theoretic approaches, such as Stackelberg game [53], are used to achieve the equilibrium state (e.g. Nash equilibrium [54]) and it involves PUs and SUs as players of the game. Examples of the non-game theoretic approaches are reinforcement learning [55] and convex optimization [56]. Table 2-1 presents the gains, functions, and characteristics of the spectrum leasing schemes. The performance enhancements achieved by each scheme is shown in Table 2-2 (see subsection 2.2.6).

Table 2-1: Gains, functions, and characteristics of the spectrum leasing schemes

Challenges	**References**	**A Gains**			**F Functions**							**C Characteristics**					
		A1 PUs Gains		A2 SUs Gains	F1 PUs Functions				F2 SUs Functions			C1 Networking topology		C2 Intra-operative mode		C3 Inter-operative mode	
		A1.1	A 1.2	A2.1	F1.1	F1.2	F1.3	F1.4	F2.1	F2.2	F2.3	C1.1	C1.2	C2.1	C2.2	C3.1	C3.2
H1 Increasing the monetary gain of PUs	[48]	×		×	×				×			×			×		×
	[57]	×		×	×						×		×		×		×
	[42]	×		×	×		×	×		×	×	×		×		×	
	[58]	×			×								×		×		×
	[45]	×		×	×						×		×		×		×
	[36]	×		×	×							×		×			×
H2 Selecting an optimal channel with white spaces by SUs	[49]		×	×		×	×				×	×			×	×	
	[52]		×	×			×	×			×	×		×		×	
	[39]		×				×	×			×	×		×		×	
	[37]		×	×			×	×			×	×		×		×	
													×		×	×	
	[59]		×	×			×	×			×		×	×		×	
	[60]		×	×			×	×			×		×	×		×	
H3 Scheduling the channel access of PUs and SUs	[61]		×	×		×	×				×		×		×		×
	[50]		×	×	×	×	×				×	×		×		×	
	[62]		×	×	×	×		×		×	×	×		×		×	
	[63]		×	×		×	×			×			×		×	×	
	[64]		×	×		×	×				×		×		×		×
	[65]		×	×		×	×				×		×		×	×	
	[66]			×					×	×			×		×		×
	[67]		×	×		×	×	×		×		×		×		×	
	[68]		×	×		×	×	×			×	×			×	×	
	[43]	×		×	×					×	×		×	×		×	
H4 Continuous monitoring of white spaces being leased to SUs by PUs	[38]			×			×				×	×			×		×
	[51]			×			×				×	×			×		×
	[69]			×			×				×	×			×		×
	[47]	×		×	×					×			×	×			×

2.2.5.1 Increasing the monetary gain of PUs

There are six spectrum leasing schemes that focus on addressing the challenge of increasing the monetary gain of PUs that motivates the PUs to participate in spectrum leasing. These schemes have been shown to increase the monetary gain of PUs, as well as to enhance PUs' or SUs' QoS performance (e.g. throughput).

2.2.5.1.1 Schemes that uses Game theoretic approaches

In [48], Alptekin and Bener (2009) propose one PU F(1) and one SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), as well as collaborator selection (F2.1) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' profit as seller in terms of its utility function U_p, which helps to satisfy the QoS parameters (e.g. jitter) of SUs as buyers, in the presence of J PUs and K SUs. The functionalities are modeled and solved using game theory and the Nash equilibrium in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. The cumulative utility function of J PUs is defined as:

$$U_p = \sum_{j=1}^{J} p_{jk}.d_{jk} - c_{jk}.d_{jk} \tag{2.1}$$

where $j = \{1,2,\ldots,J\}$, p_{jk} is the price that PU j imposes on SU k, d_{jk} is the demand factor (i.e. SU k's expectation on QoS requirement including jitter and throughput from PU j), and c_{jk} is the cost associated with the channel leased to SU k which must be paid by PU j to regulatory authorities (e.g. Federal Communications Commission (FCC)). The PU j determines the cost of white spaces (F1.1) and on that basis selects SU k if the difference between price p_{jk} and cost c_{jk} in PU utility function is positive, which indicates a monetary gain for PU j. The SU k selects a PU collaborator (F2.1) to achieve its QoS level as indicated in the demand factor d_{jk} while paying the PU j at the specified price p_{jk}. It has been shown that PUs are more likely to fulfill the SUs' QoS demand with the increment of price p_{jk} (i.e. monetary gain).

In [57], Lin and Fang (2008) propose one PU F(1) and one SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), as well as SUs' packet transmission (F2.3) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s respectively, while taking into account the mutual benefits of PUs (or sellers) and SUs (or buyers). The functionalities are modeled in the presence of J PUs and K SUs; and solved using a two-level game that is split into PU-level game and SU-level game in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. In this hierarchy of games; PUs compete with each other to lease their spectrum to SUs by adjusting their price of white spaces in order to maximize their respective utility functions; each SU attempts to lease a certain amount of white spaces from PU that provides the optimal quality white spaces. The PUs' $j \in J$ utility function is defined as:

$$U_{p,j} = \sum_{k=1}^{K} B_{jk}\{p_{jk} - c_j\} \tag{2.2}$$

where B_{jk} is the bandwidth (or white spaces) that PU j allocates to SU k, p_{jk} is the price that PU j imposes on SU k, and c_j is the cost associated with the channel leased to SU k which must be paid by PU j to regulatory authorities (e.g. FCC). A PU decides to play a game if price p_{jk} is greater than cost c_j of the leased channel (F1.1). The SUs' utility function is defined as:

$$U_s = \begin{cases} log_2\left(1 + R_{s,k}\right) & Case - I \\ log_2\left(1 + {R_{s,k}}^{MAX}\right) & Case - II \end{cases} \tag{2.3}$$

where $\mathrm{R_{s,k}}$ and ${\mathrm{R_{s,k}}}^{MAX}$ are the transmission rate of SU k. In Case-I PU allocates lesser white spaces to SU k than it demands; while in Case-II PU allocates higher bandwidth to SU k than it demands. The higher the amount of white spaces provided by PU to SU, the more is the transmission rate of SU k (F2.3). It has been revealed that, the number of SUs increases with the price of white spaces that PUs impose to SU.

In [42], Yi et al. (2010) propose three PU F(1) and two SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), relay selection (F1.3) and PUs' packet transmission (F1.4), as well as determination of SU's channel access time (F2.2) and SUs' packet transmission (F2.3) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' and SUs' network utility functions, U_p and U_s, respectively. The PUs and SUs are rational and selfish in nature. The functionalities are modeled and solved using Stackelberg game, in which the PU is the leader and the SU is the follower in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The Nash equilibrium maximizes both PUs' and SUs' utility functions, U_p and U_s. The PUs' utility function is defined as:

$$U_p = u_d + u_r \tag{2.4}$$

where u_d and u_r are revenues. Revenue u_d is dependent on the ratio of total PUs' packet transmissions, which include successful packet transmissions through direct transmissions (i.e. from PU host to PU BS) and relaying through SUs, to total traffic demand of all PU hosts. Revenue u_r is derived from the white spaces being leased to SUs. The SUs' utility function is defined as:

$$U_s = u_s - u_r \tag{2.5}$$

where u_s is derived from the total SUs' packet transmissions from all SU hosts. Both U_p and U_s take into account the SNR of the channels. There are two main steps in the Stackelberg game. Firstly, the PU BS (or leader) determines its strategy comprised of a set of potential SU relaying nodes (F1.3) and the costs (i.e. the price of white spaces per unit access time) to be imposed on SUs (F1.1), and sends the PUs' strategy to SU BS. Using the fixed leader's strategy, the SU BS (or follower) determines the amount of white spaces to request from PUs based on the costs (F2.1); hence, higher cost may reduce the amount of white spaces to request. The SU BS sends the SU strategy to PU BS. Secondly, using a fixed follower's strategy, the PU BS selects relay nodes and finalizes the costs, and starts packet transmissions (F1.4). Similarly, the SU BS allocates the leased white spaces amongst SUs for their respective packet transmission (F2.3). The spectrum leasing scheme has been shown to increase PUs' and SUs' utility functions, U_p and U_s, as well as to increase the amount of white spaces being leased. This scheme also decreases the price of white spaces per unit access time.

2.2.5.1.2 Schemes that uses non-Game theoretic approaches

In [58], Kim and Shin (2009) propose one PU F(1) function, namely determination of the cost of white spaces (F1.1) in order to increase PUs' monetary gain (A1.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' profit by controlling the SUs' admission and eviction strategies. The admission strategy allows the SUs to utilize PUs' channels on the basis of the requested amount of white spaces, which basically yields the PUs' profit. Hence, if SUs demands a small amount of white spaces, then PUs may reject their admissions due to the less monetary gain. This is because the PUs are interested to allocate white spaces to SUs that request larger amount of white spaces in order to maximize their monetary gain. Whereas, the eviction strategy is set so that SUs evacuate the channel immediately if PUs' activities reappear. The function is modeled and solved using semi-Markov decision process and linear programming in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. The PUs allocates their underutilized channels to a group of k SUs. The expected revenue of PUs is defined as:

$$r_P = \sum_{k \in K} p_k \, Q_k K \tag{2.6}$$

where p_k is the price that k SUs pay to PU in return of its QoS demand Q_k, while K is the number of SUs in the group. Higher PUs' revenue, which comes with higher price of white spaces (F1.1) indicates higher QoS demand from SUs. It has been shown that PUs' revenue increases with the amount of white spaces. However, the PUs' revenue decreases when the white spaces become oversupplied.

In [45], Song and Lin (2009) propose one PU F(1) and one SU functionalities, namely determination of the cost of white spaces (F1.1), as well as SUs' packet transmission (F2.3) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the profit of PUs while allocating the white spaces to SUs. The function is modeled and solved using auction-based property-rights model mechanism in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. In a property-rights model, SUs are divided into non-overlapping groups and a leader is elected from each group. The auction mechanism is divided into time windows, and each window is further divided into two phases, namely auction and communication. There are four main purposes in regards to the auction mechanism. Firstly, it maximizes the overall spectrum utilization. Secondly, it maximizes the number of SU winners (or SU groups that gain a channel). Thirdly, it fulfills the bandwidth requirement of SUs. Note that, the channels are heterogeneous and each channel has different amount of bandwidth (or white spaces). Fourthly, it maximizes the PUs' revenue. In a round of bidding, each SU leader determines a bid value based on hungry degree, which takes into account the amount of white spaces required by its group of SUs. During the auction phase, the PU auctions off n channels with white spaces to m SU leaders in two phases. Each SU leader uses an auction phase, which is based on its bandwidth requirement, to bid for a leasing channel. Higher value of hungry degree leads to higher bid value. During the first phase of auction, in order to meet the first, second and third purposes, the PU grants channels to as many groups of SUs as possible to meet their respective minimum requirement on the amount of white spaces. During the second phase of auction, in order to achieve the fourth purpose, the PU allocates the channels with white spaces to SU leaders that offer higher bid values (F1.1). During the communication phase (F2.3), the SUs transmit packets, and the PU keeps track of available white spaces for auctions in the next time window. The spectrum leasing scheme has been shown to increase throughput performance in regards to vacant channels.

In [36], Wu et al. (2008) propose one PU F(1) function, namely determination of the cost of white spaces (F1.1) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PU monetary gain and SUs network utility function U_s,, while preventing the collusive SUs to access the PUs' white spaces. The collusive SUs form a coalition and deliberately decrease the price of white spaces offered by PUs. The function is modeled and solved using binary linear programming and convex optimization in an intra-cooperative (C2.1) mode and non-intercooperative (C3.2) mode, respectively. Binary linear programming is a mathematical method to determine the optimal results that comprises on binary integers (i.e. 0 and 1). The PU sells white spaces to K SUs with the assistance from a third-party spectrum broker. Upon the reception of bid values $b_k = \{b_1, b_2, \ldots, b_K\}$ from K SU, the spectrum broker announces the winning SUs by defining the channel allocation $x_k = \{x_1, x_2, \ldots, x_K\}$ and the associated price $p = \{p_1, p_2, \ldots, p_K\}$ for K SUs. For the winning SUs, the channel allocation x_k set to one (i.e. $x_k = 1$), which indicates that the channel has been allocated to winner SU k. The gain of each winning SUs' is g_k, which lead to an efficient channel allocation which is used to compute the utility function of SUs i.e. $U_s = \sum_{k=1}^{K} g_k . x_k$. Higher values of U_s indicate higher number of winning SUs in the auction for white spaces. It has been shown that, as the number of winning SUs increases the price of the white spaces imposed by the PUs as sellers also increases.

2.2.5.2 Selecting an optimal channel with white spaces by SUs

There are six spectrum leasing schemes that focus on motivating the SUs to participate in spectrum leasing by increasing the amount of white spaces for SUs. These schemes have been shown to enhance PUs' or SUs' QoS performance (e.g. throughput).

2.2.5.2.1 Schemes that uses Game theoretic approaches

In [49], Chan et al. (2011) propose two PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time (F1.2) and relay selection (F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the spectrum utilization of PU and SU networks by adopting the cooperation strategies in between of J PUs and K SUs in the form of PUs and SUs utility functions, U_P and U_S, respectively. In separate cooperation, PU j and SU k form a one-to-one collaborative relationship with each other, while in grand cooperation, PUs and SUs form a coalition that comprises of many one-to-one and one-to-many collaborative relationships with each other. The functionalities are modeled and solved using canonical coalition game theoretic framework and convex optimization problem in a non-intracooperative (C2.2) mode and inter-cooperative (C3.1) mode, respectively. The PU utility function is defined as:

$$U_P = u(R_P) + p_{SP_j} - L(c_P) \tag{2.7}$$

where $u(.)$ and $L(.)$ are concave function that maps the PU achievable transmission rate R_P as utility gain and PU cost c_P as utility loss; while p_{SP_j} is the price of white spaces that PU j imposes on SUs. The SU utility function is defined as:

$$U_S = r(R_S) + P_{S_kP} \tag{2.8}$$

where $r(.)$ is concave function that projects SU achievable rate R_s as revenue and P_{S_kP} is the price that PUs imposes on SU k in order to lease its channel. It has been shown that, the grand cooperation strategy produces higher optimal utility value than individuals' cooperation.

In [52], Vilar et al. (2010 propose two PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3), and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s in the presence of a PU communication node pair in order to minimize the SUs' interference to PUs by reducing their power consumption. The PU determines the maximum allowable interference that PU can tolerate from SUs I_{max}^P; while the SUs aim to reduce their transmission power in order to fulfill the requirement I_{max}^P. The function is modeled and solved using Stackelberg game in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. In this scheme the PU is the leader and the SU is the follower. To foster collaboration with SUs, the PU maximizes its utility function, and it is defined as:

$$U_p = u_p(I_k^S, \Delta T_P^p) \tag{2.9}$$

where u_p increases with the increment of interference from SU k (or $I_k^S \leq I_{max}^P$) and decreases with the increment of PUs' transmission power ΔT_P^p. To foster collaboration with PU, the SU maximizes its utility function, and it is defined as:

$$U_s = u_s(R_k^S, I_k^S) \tag{2.10}$$

where u_s increases with the increment of the SU transmission rate R_K^S and decreases with interference from SU I_k^S. Note that, R_k^S and I_k^S are increases with the SU transmission power p_k, and hence $R_k^S(p_k)$ and $I_k^S(p_k)$. Maximizing U_s helps to maximize the SU k's power vector $P_K^{I_{max}^P} = \text{argmax}_p\{u_s(R_k^S, I_k^S)\}$. This has led to computing the overall utility function of PUs and SUs on the basis of I_{max}^P. The PU selects a SU relay node $k \in K$ (F1.3) that has the lowest transmission power for transmission of PU packets (F1.4), as well as SUs' packets (F2.3) among the other SUs. It has been shown that the proposed scheme achieves higher utility function for both PUs and SUs compared to the traditional scheme.

2.2.5.2.2 Schemes that uses non-Game theoretic approaches

In [39], Cao et al. (2012) propose two PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3) and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) in centralized (C1.1) SU networks. The purpose is to maximize the spectrum utilization of PU and SU networks, where the PU and SU BSs operate in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The PU source node i selects the best available SU relay node k, and establishes communication with the PU destination node j. The SU relay is used to transmit PU and SU packets using a quadrature modulation scheme, which depends on two factors, namely power allocation factor $0 \leq F_{ij,k}^p \leq 1$ and weight factor $0 \leq w_{ij,k}^p \leq 1$. The power allocation factor determines the transmission of packets through SU relay node. Note that, the SU relay node transmits PU packets only if $F_{ij,k}^p = 1$, the SU packets only if $F_{ij,k}^p = 0$, and both PUs' and SUs' packets if $0 < F_{ij,k}^p <$

1. Whereas, the weight factor determines the respective throughputs of PU and SU network, respectively. The selected SU relay node k transmits PU and SU packets simultaneously using transmission power $\mathrm{P}^{s}_{\mathrm{ij,k}}$ in two orthogonal channels (i.e. in-phase and quadrature channels) exploited using a quadrature modulation approach. The SU relay node relays PU packets using transmission power $F^{p}_{ij,k} \cdot P^{s}_{ij,k}$ using in-phase channel, and sends SU packets using transmission power $(1 - F^{p}_{ij,k}) \cdot P^{s}_{ij,k}$ in quadrature channel. The throughput of PUs and SUs is represented by a weighted sum throughput T_T, which is defined as:

$$T_T = (1 - w^{p}_{ij,k}) \cdot T_p + w^{p}_{ij,k} \cdot T_s \tag{2.11}$$

where T_p and T_s represent PUs' and SUs' throughput, respectively. Note that, $T_T = T_p$ if $w^{p}_{ij,k} = 0$, and $T_T = T_s$ if $w^{p}_{ij,k} = 1$; while T_p and T_s achieve a balance if $w^{p}_{ij,k} = 1/2$. A primal-dual subgradient algorithm, including Lagrange multipliers and the Karush-Kuhn-Tucker conditions, is used to optimize $F^{p}_{ij,k}$ and $P^{s}_{ij,k}$ in order to optimize the weighted sum throughput T_T. The PU selects a SU only if it improves throughput performance (F1.3); while the selected SU transmits the PU and SU packets simultaneously (F1.4), or the SU packets only (F2.3) when the PU is inactive. Through achieving a balanced throughputs T_p and T_s, the scheme has been shown to maximize T_T, and this is due to the dependence of T_p and T_s on power allocation factor $F^{p}_{ij,k}$ and weight factor $w^{p}_{ij,k}$.

In [37], Jayaweera et al. (2011) propose two PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3) and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs

(A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) and distributed (C1.2) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s respectively; in terms of power savings of PUs when it collaborates with SUs in the presence of J PUs and K SUs. For centralized CRNs, the functionalities are modeled and solved using reinforcement learning in an intra-cooperative (C2.1) mode, and inter-cooperative (C3.1) mode, respectively. Whereas for distributed CRNs, the functionalities are modeled and solved using reinforcement learning in a non-intracooperative (C2.2) mode and inter-cooperative (C3.1) mode, respectively. The PU $j \in J$ utility function is defined as:

$$U_p = \frac{P_{j,d} - P_j\left(P_{k(j),j}\right)}{P_{j,d}}(R_j(\alpha_j) - R_{j,min}) \tag{2.12}$$

where $P_{j,d}$ is the maximum transmission power of PU j through direct PU-PU transmission without using a SU relay node, $P_j\left(P_{k(j),j}\right)$ is the PU j transmission power through PU-SU-PU transmission using SU k as a relay node where $P_{k(j)}$ is the transmission power for SU k to relay the PUs' packets to its destination, $R_j(\alpha_j)$ and $R_{j,min}$ are the achievable transmission rate of PU j after allocating α_j of white spaces to SUs and the minimum transmission rate of PU j for direct transmission, respectively. The SU $k \in K$ utility function is defined as:

$$U_{s,k} = \alpha_j W_j \log(1 + SNR_{k,i})(BER_{kj,min} - BER_{kj,(P_{k(j)})}) \tag{2.13}$$

where W_j is the bandwidth used by SU to transmit its own signal, $SNR_{k,i}$ is the signal-to-noise ratio of SU k, while $BER_{kj,min}$ and $BER_{kj,(P_{k(j)})}$ are the minimum and observed Bit

Error Rate (BER) values of SU k while relaying PU $j's$ packets. It has been shown that the transmission power of PU decreases with increasing the transmission power of SU.

In [59], Murawski and Ekici (2011) propose two PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3) and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the throughput of PUs and SUs in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The network considers a single PU source node that communicates with a PU destination node through direct PU-PU transmission or indirect PU-SU-PU transmission via SU relay node. The PU destination node transmits Request to Send (RTS); while the SU replies with Request to Cooperate (RTC) composed of channel state information upon receiving RTS from the PU. Subsequently, the PU destination node selects the suitable SUs as relay nodes using the channel state information. The criterion adapts by PU for suitable SU relaying node is on the basis of higher throughput value of a given PU-SU-PU link with respect to throughput value of PU-PU direct link. The PU destination node respond clear to coordinate (CTC) message to selected SU relay node, which indicates that given PU-SU-PU link offers higher throughput than the PU-PU direct link; whereas, if the throughput offers by PU-SU-PU link is lower than the PU-PU direct link then the PU destination node respond clear to send (CTS) message to SU relay node, which indicates the direct link PU-PU communication takes place. For the calculation of expected throughput value either from PU-SU-PU link or from PU-PU direct link, a backoff mechanism of distributed coordination function [70] is used. The expected throughput value is dependent on the probability of successful packet transmission P_s, packet transmission time t_{packet}, collision detection time $t_{collide}$, and the expected size of PU packets $E_{packet\ size}$. Furthermore, for attaining a higher throughput gain, adaptive modulation schemes (e.g. BPSK, QPSK and 16-QAM) is used with respect to the SNR of the channels. It has been shown that, the higher the throughput value can be achieved with the change of adaptive modulation scheme from BPSK to QPSK and from QPSK to 16-QAM; and the increase in number of SUs as relaying node decreases the throughput of PUs significantly due to the communication overheads.

In [60], Toroujeni et al. (2012) propose two PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3), and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to increase the link reliability by maximizing the transmission rate of a PU communication node pair and K SUs. The functionalities are modeled and solved using Orthogonal Frequency Division Multiplexing (OFDM) [71] symbols in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. There are a total of $N_S + N_{PP}$ OFDM symbols, in which N_{PP} symbols are dedicated for a PU-PU communication node pair for direct transmission, and the N_S symbols are dedicated for PU-SU and SU-SU transmissions, respectively. The PU selects the maximum transmission link R_P either from PU-PU direct link R_{pp} or from PU-SU-PU relayed link R_{psp}, and it is defined as:

$$R_P = \max\{R_{pp}, R_{psp}\} \tag{2.14}$$

Each SU $k \in K$ chooses the best channel to relay the packets from PU source node to PU destination node as well as its own packets to another SU. The SU cooperates with

PU if SU-SU transmission rate R_{ss} is equal to the price p_k charged by PU times the SU-PU transmission rate R_{sp}, and it is defined as:

$$R_{ss} = \sum_{k=1}^{K} p_k \,.\, R_{sp} \tag{2.15}$$

The higher value of R_{ss} indicates higher achievable transmission rate between SU relay node and PU destination node. It has been shown that, as the distance increases between PU source node and SUs, it decreases the number of selected SUs as relaying nodes. Furthermore, higher cost being incurred by SUs reduces the achievable transmission rates of PUs although it increases the achievable transmission rates of SUs.

2.2.5.3 Scheduling the channel access of PUs and SUs

There are ten spectrum leasing schemes that focus on scheduling of channel access time in between of PUs and SUs for their respective transmission. These schemes have been shown to enhance PUs' and SUs' QoS performance (e.g. throughput).

2.2.5.3.1 Schemes that uses Game theoretic approaches

In [61], Chen et al. (2011) propose two PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time (F1.2) and relay selection (F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' and SUs' network utility functions U_p and U_s in the presence of J PUs and K SUs. The functionalities are modeled and solved using a three-tier game in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. The PU and SU network communicate with each other using a control channel protocol in order to participate and achieve a game equilibrium Both PUs and SUs are rational in nature. The PU selects the suitable SUs as relay nodes to transmit PU's packets in order to increase its transmission rate; and the SUs in return achieve a portion of channel access time set by the PU to maximize their transmission rate. The PU divides the transmission period into three phases. The first phase is for primary transmission (PU-PU and PU-SU) during which the PUs transmit their packets to other PUs and SUs. The second phase is for relayed transmission (SU-PU) during which the SUs help the PUs to relay PUs' packets. Whereas, the third phase is for secondary transmission (SU-SU) during which the SUs transmit their own packets. The length of the primary transmission phase is α, the relay nodes transmission phase is $(1-\alpha)(1-\beta)$, and the secondary transmission phase is $(1-\alpha)\beta$. Higher value of α indicates that PUs is willing to lease its spectrum to SUs; while higher value of β encourages SUs to collaborate and relay PUs' packets. Thus, the PU must determine optimal values of α and β (F1.2) that maximize its own and SUs' transmission rate. The PU j utility function is defined as:

$$U_{p,j} = min\{\alpha R_{PS,k}, (1-\alpha)(1-\beta)R_{SP,k}\} \tag{2.16}$$

where R_{PS} and R_{SP} are the maximum transmission rate through SU relay nodes (F1.3). The SU k utility function is defined as:

$$U_s = \beta R_{SS} - (1-\alpha)p_s P_s \tag{2.17}$$

where p_s is the cost of per unit power P_s consumed by SU k as relay node to transmit PU source node packet to PU destination node. Therefore, the utility function of SU k is the difference between its revenue in terms of achievable rate R_{SS} (F2.3) and the cost of power which SU k must bear in order to relay the PU's packets. It has been shown that as the distance increases between the PU and SUs, their utility functions increases until a certain limit which then decrease.

In [50], Huang et al. (2011) propose three PU F(1) and one SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), determination of PUs' and SUs' channel access time(F1.2) and relay selection (F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s in the presence of J PUs and K SUs. The functionalities are modeled and solved using canonical coalition game in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The PU divides a unit time slot into three sub-slots for primary transmission (PU-PU and PU-SU), relayed transmission (SU-PU), and secondary transmission (SU-SU), respectively. The length of the primary transmission sub-slot is $1-\alpha$, the relay nodes transmission sub-slot is β, and the secondary transmission sub-slot is $\alpha-\beta$. Higher value of α indicates that PUs are willing to lease its spectrum to SUs; while higher value of β encourages SUs to collaborate more and relay PU packets. Thus, the PU must determine the optimal values of α and β that maximize its own as well as SUs' transmission rate. The PU $j \in J$ utility function is $U_p = F(R_p)$, where $F(.)$ is concave an increasing function that represents PUs' gain and R_p is the minimum achievable transmission rate, which can be either from PU-SU or from SU-PU, and dependent on transmitter power P_t, channel gain G and noise level σ^2. The SUs' utility function is $U_s = G(R_s) - p_s$, where G(.) is concave an increasing function that represents SUs' gain and p_s is the price that SU need to pay in order to lease channels from PUs. It has been shown that as the SUs' channel access time increases, the transmission rate of SUs increases significantly, which increases the PUs monetary gain while decreases its transmission rate since SUs uses more power to transmits its own packets.

In [62], Lei et al. (2010) propose three PU F(1) and two SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), determination of PUs' and SUs' channel access time(F1.2), and relay selection (F1.3), as well as determination of SU's channel access time (F2.2) and SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s, respectively, in the presence of a PU communication node pair and K SUs. The functionalities are modeled and solved using Stackelberg game in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The PU divides the transmission period into three phases. The first phase is for primary transmission (PU-PU and PU-SU) during which the PUs transmit their packets to other PUs and SUs. The second phase is for relayed transmission (SU-PU) during which the SUs help the PUs to relay PUs' packets. Whereas, the third phase is for secondary transmission (SU-SU) during which the SUs transmit their own packets. The length of the primary transmission phase is $\frac{T-t_S}{2}$, the relay nodes transmission phase is $\frac{T-t_S}{2}$, and the secondary transmission phase is t_S. The PU utility function is defined as:

$$U_p = G_{SNR}(SNR_{PP} + SNR_{PSP})\frac{T-t_S}{2T} \quad (2.18)$$

where G_{SNR} is the channel gain per unit SNR, SNR_{PP} and SNR_{PSP} are the SNR values of PU-PU direct link and PU-SU-PU relayed link. Whereas, the SUs' utility function is defined as:

$$U_S = G_{t_S} - c\frac{T-t_S}{2}\{\frac{(SNR_{PS}+1)p.t_S.\sigma^2}{(SNR_{PS}-pt_S)G_{PP}}\} \quad (2.19)$$

where c is the cost per unit energy consumption, p is the price that SUs needs to bear in order to buy white spaces from PUs and σ^2 is the noise variance. It has been shown that as the distance increases between the PU and SUs, their utility functions increases until a certain limit which then decrease.

In [63], Stanojev et al. (2008) propose two PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time (F1.2) and relay selection (F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' transmission rate and the SUs' utility function. The PU divides a unit time slot into three sub-slots for primary transmission (PU-PU and PU-SU), relayed transmission (SU-PU), and secondary transmission (SU-SU), respectively. The length of the primary transmission sub-slot is $(1-\alpha)\cdot t_{slot}$, the relay nodes transmission sub-slot is $\alpha \cdot \beta \cdot t_{slot}$, and the secondary transmission sub-slot is $\alpha \cdot (1-\beta) \cdot t_{slot}$. Higher value of α and lower value of β encourages SUs to collaborate, and so the PU must determine optimal values of α and β, while maximizing its own transmission rate. The functionalities are modeled and solved using Stackelberg game in a non-intracooperative (C2.2) mode and inter-cooperative (C3.1) mode, respectively. In this scheme PU is the leader and SU is the follower. The game aims to foster collaboration between PUs and SUs by maximizing the PUs' transmission rate and enhancing the SUs' utility function. The PU source node i chooses a set of SU relay node k that provides an optimum value of PU transmission rate, which is dependent on the transmission rate from PU source node i to SU relaying node k, or $R_{ij,k}^{ps}$; while SU relaying node k calculates the transmission rate from SU relay node k to PU destination node j, or $R_{ij,k}^{sp}$; as well as β. Hence, the value of β must be chosen carefully to encourage collaboration between PU and SU. The choice of β must maximize the SU-PU transmission rate $(\alpha \cdot \beta \cdot t_{slot}) \cdot R_{ij,k}^{sp}$; on the other hand, the choice of α must maximize the SU-SU transmission rate $\{(\alpha \cdot (1-\beta)) \cdot t_{slot}\} \cdot R_{k}^{ss}$. The optimal value of β is $\hat{\beta} = argmax_{\beta \in [0,1]} \beta . R_{ij,k}^{sp}$; and $\hat{\beta}$ is applied in the calculation of $\hat{\alpha} = f(1/\hat{\beta})$. The PU source node selects a suitable SU relay node (F1.3) to transfer its packets to PU destination node (F1.4) if SU relay node provides higher transmission rate, otherwise it chooses PU-PU direct link. The PU calculates channel access time for PUs' and SUs' (F1.2). It has been shown that, as the number of SU relay nodes increases, the outage probability of PU decreases and the transmission rate of SUs increases. The SUs aim to maximize their utility function in order to transmit its own packets (F2.3). The SUs utility function is $u_{P_k,\mathrm{P}_{-\mathrm{k}}}^{SS}$, where P_k is the transmission power of SU relaying node k, $\mathrm{P}_{-\mathrm{k}}$ is a vector of the transmission power of the SU non-relaying nodes. The PU adjusts $\hat{\beta}$ to determine the time distribution among PUs' and SUs' (F1.2) transmissions, and this is followed by the selection of the best available SUs as relay nodes (F1.3) for possible communication between a PU node pair. It has been shown that the PUs' and SUs' throughput performances can be increased by increasing the number of SU relay nodes k and decreasing the distance between PU and SU.

In [64], Wang et al. (2010) propose two PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time (F1.2) and relay selection

(F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s in the presence of a PU communication node pair and K SUs. The functionalities are modeled and solved using the game theoretic approach and the Stackelberg equilibrium in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. In this game theoretic approach, PUs and SUs are rational in nature, in which the PUs and SUs attempt to achieve their respective equilibrium point. The PU selects suitable SUs that transmit PU packets as relay using their respective transmission power, while the SUs in return achieve a portion of channel access time set by the PU to transmit their own packets. The PU divides a unit time slot into two sub-slots for primary transmission (PU-PU, PU-SU, SU-PU), and secondary transmission (SU-SU), respectively. The length of the primary transmission sub-slot is α, while the secondary transmission sub-slot is $1 - \alpha$. Higher value of α indicates that PUs are willing to lease its spectrum to SUs in order to maximize its packet transmission; while allocating the remaining time to SUs for their own packet transmission. Thus, PU must determine the optimal value of α (F1.2) that maximize its own and SUs' transmission rate. The PU utility function is defined as:

$$U_p = \alpha R_p(\alpha) \tag{2.20}$$

where $R_p(\alpha)$ is the achievable transmission rate through SU relay nodes (F1.3) and it is dependent on transmitter power P_t, channel gain G and noise variance σ^2. The SUs' utility function is defined as:

$$U_s = r_k(R_k)t_k - \frac{1}{2}\alpha P_k \tag{2.21}$$

where r_k, R_k and t_k are the revenue, achievable transmission rate and allocation time of SU k, and P_k is the transmission power used by SU k to relay the PUs' packets to PU destination and therefore it is considered as a cost by SU k. Therefore, the utility function of SU k is the difference between its revenue in terms of achievable transmission rate (F2.3) and the energy cost that SU k must bear to relay the PUs' packets. It has been shown that PUs' utility function increases with the increment of the α value. Furthermore, as the distance between PUs and SUs decreases, it increases their utility functions significantly because of higher channel gain.

In [65], Zhang et al. (2010) propose two PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time(F1.2), relay selection (F1.3), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s in order to enhance their transmission rate in the presence of a PU communication node pair and K SUs. The functionalities are modeled and solved using game theory and the Nash equilibrium in a non-intracooperative (C2.2) mode and inter-cooperative (C3.1) mode, respectively. In this game, the PU selects the suitable SUs as relay nodes to transmit PUs' packets using their respective transmission power; and in return, the SUs receive a portion of channel access time set by the PU to transmit their own packets. The PU divides a unit time slot into three sub-slots for primary transmission (PU-PU and PU-SU), relayed transmission (SU-PU), and secondary transmission (SU-SU), respectively. The length of the primary transmission sub-slot is $1-\alpha$, the relay nodes transmission sub-slot is $\alpha\beta$, and the secondary transmission sub-slot is $\alpha(1-\beta)$. Higher value of α indicates that PUs are willing to lease its white spaces to SUs; while higher value of β encourages SUs to collaborate more and relay PU packets. Thus, the PU must determine optimal values of α and β (F1.2) that maximize its own and SUs' transmission rate. The PUs' utility function is defined as:

$$U_p = R_{PSP} - R_{PP} + \alpha c_p P_p \tag{2.22}$$

where R_{PSP} and R_{PP} are the achievable transmission rate through SU relay nodes (F1.3) and PU-PU direct transmission. These rates are dependent on transmission power P channel gain G and noise power N; whereas, c_p is the cost per unit of transmission power consumed by PU source node to transmit its packets to SUs and PU destination node. The SUs' utility function is defined as:

$$U_s = \alpha(1-\beta)log_2\left(1+\frac{P_s G_s}{N}\right) - \alpha c_s P_s \tag{2.23}$$

where c_s is the cost per unit transmission power consumed by SU relay node k to transmit PU source node's packets to PU destination node. Therefore, the utility function of SU k is the difference between its revenue in terms of the achievable rate (F2.3) and the energy cost that SU k must bear to relay the PUs' packets. It has been shown that, as the distance increases between the PU and SUs, their utility functions increases until a certain limit which then decrease.

In [66], Zhu et al. (2012) propose two SU F(2) functions, namely collaborative selection (F2.1) and determination of SU's channel access time (F2.2) in order to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. There are two types of markets, namely primary market (comprised of SU service providers and PUs) and secondary market (comprised of SU service providers and SU hosts). The functionalities are modeled and solved using a hierarchical game theoretic framework comprised of upper- and lower-level games and in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. The purpose is to maximize the SUs' service provider and SU network utility functions, $U_{P,i}(t)$ and $U_{s,i}(t)$, respectively. The hierarchical game theoretic framework is as follows:

- Secondary market allows SU hosts to purchase white spaces from SU service providers on a short-term basis (e.g. minutes), and it is a lower-level game modeled by evolutionary game. Each SU service provider i offers white spaces, which are represented by bandwidth b_i and price p_i. Note that, higher price p_i for a particular bandwidth b_i reduces demand levels, and so it improves network performance. Subsequently, each SU host competes and selects a SU service provider. Hence, the secondary market implements collaborator selection (F2.1). Each SU aims to maximize its individual utility function defined as:

  $$U_{s,i}(t) = \alpha \cdot b_i(t)/p_i \tag{2.24}$$

  where α is a constant based on network performance requirement, in order to maximize its network performance satisfaction. The number of SUs that choose service provider i is represented by $n_i(t)$.

- Primary market allows SU service providers to purchase white spaces from PUs (or spectrum brokers) on a long-term basis (e.g. weeks or months), and it is a upper-level game modeled by differential game. Each SU service provider i

purchases some amount of white spaces $c_i(t)$ from PUs based on the selection of SU service providers $x_i(t)$ in order to maximize profits. Hence, it implements the determination of SU's channel access time (F2.2). Note that, higher amount of the purchased white spaces improves network performance and so it attracts more SUs; however, it reduces monetary revenues. Each SU service provider i adjusts the amount of white spaces $c_i(t)$, and maximizes its profit defined as:

$$U_{P,i}(t) = p_i \cdot n_i(t) - \beta_i \cdot c_i^2(t) \tag{2.25}$$

where $p_i \cdot n_i(t)$ represents the monetary revenue, $\beta_i \cdot c_i^2(t)$ represents the cost paid to the PUs, and β_i is a constant weight. Note that, with $c_i^2(t)$, it causes the cost to increase rapidly, and so it prevents a SU service provider i from being too aggressive. At Nash equilibrium, each SU service provider obtains maximized profit. In differential game, the SU service providers make decision simultaneously; however, some providers may make decision first, and they are called the leaders. In this case, a Stackelberg differential game can be applied to achieve Stackelberg equilibrium. In Stackelberg game, the leader providers make decisions first, followed by follower providers. So, the leader providers can achieve higher payoff, and the follower providers make decision based on the optimal strategies made by the leader providers. The spectrum leasing scheme has been shown to increase SU service providers' profits.

2.2.5.3.2 Schemes that uses non-Game theoretic approaches

In [67], Asaduzzaman et al. (2012) propose three PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time(F1.2), relay selection (F1.3) and PUs' packet transmission (F1.4), as well as determination of SU's channel access time

(F2.2) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to minimize the outage probability of PUs' network and to maximize the outage capacity of SUs' network. The outage probability indicates the halt of PUs' packet transmission for a certain period of time when the transmission signal power is less than a certain threshold value; while the outage capacity is the SUs' transmission rate during outage. Hence, generally speaking, the functionalities are based on transmission rate and channel access duration of PUs and SUs in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The network considers a PU communication node pair, and it is separated by a single centralized SU network comprised of potential SU relaying nodes K. The PU source node i selects the best available SU relaying node $k \in K$, and creates a multiple-hop communication with the PU destination node j. The PU source node makes decision whether to communicate directly or through SU relaying nodes to the PU destination node. The selection of SU relaying node k is based on the transmission rate offered by itself in a PU-SU-PU communication, $R_{ij,k}^{psp}$. The $R_{ij,k}^{psp}$ is computed separately in two steps. Specifically, PU source node i calculates the transmission rate from PU source node i to SU relaying node k, or $R_{ij,k}^{ps}$; and SU relaying node k calculates the transmission rate from SU relaying node k to PU destination node j, or $R_{ij,k}^{sp}$. Subsequently, the PU source node i selects the best available SU relaying node $k \in K$ based on the transmission rate of the bottleneck link or $R_{ij,k}^{psp} = \min\{R_{ij,k}^{ps}, R_{ij,k}^{sp}\}$. The PU source node i communicates through SU relaying node k when $R_{ij,k}^{psp} > R_{ij}^{pp}$, otherwise, the PU source node i chooses to communicate directly with PU destination node, where R_{ij}^{pp} represents the transmission rate of PU-PU direct transmission. Note that, the transmission rates $R_{ij,k}^{psp}$ and R_{ij}^{pp} are dependent on SNR. The PU divides the transmission period into three phases. The first phase is for primary transmission (PU-PU and PU-SU) during which the PUs transmit their packets to other PUs and SUs. The second phase is for relayed transmission (SU-PU) during which the SUs help the PUs to relay PUs' packets. Whereas, the third phase is for secondary transmission (SU-SU) during which the SUs transmit their own packets. The first two phases, namely A and B, are allocated to the transmission of PU packets, specifically PU-SU and SU-PU, respectively; while the third phase C is for SU-SU transmission. Hence, the outage capacity of SU is dependent on the time duration of phase C and the transmission rate of SU-SU. Denote the requirement on PU's transmission rate as R_{ij}, the outage of PU occurs whenever $R_{ij,k}^{psp} < R_{ij}$ and $R_{ij}^{pp} < R_{ij}$. The PU source node selects a suitable SU relay node (F1.3) to transfer its packets to PU destination node (F1.4) if SU relaying node provides higher transmission rate, otherwise it chooses PU-PU direct link. The PU calculates channel access time for PUs (F1.2) and SUs (F2.2). It has been shown that as the number of SU relaying nodes increases, the outage probability of PU decreases and the transmission rate of SUs increases.

In [68], Khalil et al. (2011) propose three PU F(1) and one SU F(2) functionalities, namely determination of PUs' and SUs' channel access time (F1.2), relay selection (F1.3) and PUs' packet transmission (F1.4), as well as SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximize the PUs' and SUs' utility functions U_p and U_s, respectively. The functionalities are modeled and solved using Lyapunov Optimization [72] in a non-intracooperative (C2.2) mode and inter-cooperative (C3.1) mode, respectively. The PU divides a unit time slot into three sub-slots for primary transmission (PU-PU and PU-SU), relayed transmission (SU-PU), and secondary transmission (SU-SU), respectively. The length of the primary transmission sub-slot is $1-\alpha$, the relay nodes transmission sub-slot is $\alpha\beta$, and the secondary transmission sub-slot is $\alpha(1-\beta)$. The main objective of the PU j's utility function is to improve its transmission rate, and it can be computed with and without the cooperation from SUs as relay nodes as follows:

$$U_p = \begin{cases} R_{p,j} & PU - PU\ transmission \\ (1-\alpha).R_{p,jk} & PU - SU - PU\ transmission \end{cases} \tag{2.26}$$

where $R_{p,j}$ represents the PUs' achievable direct transmission rate without any cooperation with SUs and $R_{p,jk}$ represents the achievable PUs' transmission rate in cooperation with SUs as relaying nodes. Higher $U_p = R_{p,j}$ is applied when PUs' direct transmission rate is greater than the PUs' transmission rate in cooperation with SUs, otherwise $U_p = (1-\alpha)R_{p,jk}$ is applied. The PU cooperates with SU when transmission rate is at least equal to its minimum transmission rate $R_{p,j}$ (or $R_{p.j} \leq R_{p,jk}$). The main objective of SU k's utility function is to improve its own transmission rate which is defined as:

$$U_s = \alpha(1-\beta)R_{s,k} \tag{2.27}$$

where $R_{s,k}$ represents the transmission rate of SU k. Higher U_s indicates that the transmission rate of SU k increases due to the higher amount of channel access time being allocated for its own transmission. It has been shown that, the proposed scheme achieves higher transmission rate due to the cooperation between PUs and SUs.

In [43], Zhou et al. (2011) propose one PU F(1) and two SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), as well as determination of SU's channel access time (F2.2) and SUs' packet transmission (F2.3) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. The purpose is to enable the SUs to acquire the white spaces efficiently when PUs intends to lease it in order to maximize the monetary gain of PU and transmission rate SU networks. The functionalities are modeled and solved by introducing rules for spectrum management and spectrum leasing in an intra-cooperative (C2.1) mode and inter-cooperative (C3.1) mode, respectively. The spectrum management rule is set by the PU BS to regulate the spectrum leasing process in order to maximize PUs' revenue F(1.1) and guarantee a fair spectrum trade market by offering the discounted spectrum price to SUs in combination with spectrum and time optimization. The spectrum leasing rule is set by the SUs, through which SUs takes the decision to acquire the white spaces from PUs if it fulfills the bandwidth requirements desired by SUs for a specified period of time (F2.2), which SUs mentioned to PU BS for its packet transmission (F2.3). It has been shown that as PU allocates more channel bandwidth to SUs while increasing the number of transmission slots, it maximizes the SUs transmission rate and throughput.

2.2.5.4 Continuous monitoring of white spaces being leased to SUs by PUs

There are four spectrum leasing schemes that focus on the monitoring of SUs' channel access activities in spectrum leasing by PUs', so that SUs are ensued to follow (or fulfill) suit according to spectrum leasing contract with PUs. These schemes have been shown to enhance PUs' or SUs' QoS performance (e.g. throughput).

2.2.5.4.1 Schemes that uses Game theoretic approaches

In [9], Jayaweera et al. (2010) propose one PU F(1) and one SU F(2) functionalities, namely relay selection (F1.3) and SUs' packet transmission (F2.3) in order to enhance the network performance of PUs (A1.2) and to provide dedicated channel access to SUs (A2.1) in centralized (C1.1) SU networks. The purpose is to maximze the PUs' and SUs' utility functions, U_P and U_S, respectively. Both PUs and SUs are rational and selfish in nature. The functionalities are modeled and solved using a game theoretic framework in a non-intracooperative (C2.2) mode and non-intercooperative (C3.2) mode, respectively. The Nash equilibrium maximizes both PUs' and SUs' utility functions, U_P and U_S. Each PU actively adjust an interference cap I_C, which is the maximum level of interference from SUs I_{SU}. The PU selects those SUs (F1.3) that not violates the interference cap I_C, so that PU achieve its minimal SNR and QoS level. The PUs' utility function is defined as:

$$U_P = (I_{C,max} - (I_C - I_{SU}))I_C \tag{2.28}$$

The PU constantly broadcasts the I_C and I_{SU} to all SUs. Whereas, each SU adjusts its transmission power to ensure the current level of interference from SUs I_{SU} is lower than I_C, in order to maximize its own rewards in terms of higher throughput for packet transmission (F2.3). The SU utility function is defined as:

$$U_S = (I_C - \lambda_S I_{SU})r_S \tag{2.29}$$

where λ_S and r_S are positive coefficient and reward function, respectively. The spectrum leasing scheme has been shown to increase PUs' and SUs' utility functions U_P and U_S, as well as to increase the rewards (i.e. transmission rate per user). Similar schemes have also been applied in (Jayaweera and Li, 2009) [51] and (Hakim et al., 2010) [69] as follows:

- In [51], the purpose is to examine the power control mechanism and its effect on the utility function of PUs. The PUs' utility function, which aims to achieve the required QoS performance of PUs and SUs, is defined as:

$$U_P = I_C - (I_C - I_{SU})^2 - (e^{(I_{SU} - I_c)} - 1) \quad (2.30)$$

Whereas, the SU utility function, which aims to achieve SUs' energy efficiency, is defined as:

$$U_s = \frac{R_{s,k}\left(1 - e^{[0.5(SNR_s)]}\right)}{p_k} \quad (2.31)$$

where $R_{s,k}$ and p_k are the transmission rate and transmission power of SU k. The SUs' utility function defines SUs' packet transmission (F2.3), which represents the number of successful transmitted bits per unit of transmission power.

- In [69], the propose is to adjust the PUs' interference level in accordance with the SUs' transmission requirements of SNR and QoS levels, so that PUs and SUs maximize their respective utility functions. The PUs' utility function is defined as:

$$U_P = \left\{\left(I_{C,max} - (I_C - I_{SU})\right) I_C\right\}.r_P \quad (2.32)$$

where r_P is a continuous reward function that defines the PUs' gain while leasing its spectrum to SUs. The SUs' utility function is defined as:

$$U_S = \frac{r_S}{1 + e^{\lambda(I_{SU} - I_C)}} \quad (2.33)$$

where r_S is the reward function of SUs, which depends on the transmitting power of SU $k \in K$.

2.2.5.4.2 Scheme that uses non-Game theoretic approaches

In [47], Sodagari et al. (2011) propose one PU F(1) and one SU F(2) functionalities, namely determination of the cost of white spaces (F1.1), as well as determination of SU's channel access time (F2.2) in order to increase PUs' monetary gain (A1.1) and to provide dedicated channel access to SUs (A2.1) in distributed (C1.2) SU networks. Generally speaking, SUs send private information to PUs regarding their channel access time (i.e. arrival and departure times) and bid values during the auction process in which the PUs provide suitable channel allocations to SUs. There are two types of SUs, namely truthful SUs and collusive SUs. Truthful SUs provide the private information to PUs; while the collusive SUs collaborate among themselves through sharing the private information and subsequently misreport the information in order to gain the channel access. There are two approaches to misreport the information. Firstly, the collusive SUs share the bid values so that the SUs either set the bid value to the lowest or slightly higher values. Secondly, the collusive SUs share the arrival time so that the SUs either set to the arrival time to the latest or slightly earlier values, and this minimizes the competitiveness among the SUs for channel access in auctions and subsequently minimizes the bid values. The functionalities are modeled and solved using an approach called Dominant Strategy Incentive Compatible (DSIC) in which a SU can reduce its payment to the PUs in an auction process without collusion, in an intra-cooperative (C2.1) mode and non-intercooperative (C3.2) mode, respectively. Specifically, with respect to SU k, denote the bid value by $\mathrm{b_k}(\pi_k)$ and the price p_{jk} set by the PU j, the SU adopts a π_k policy to determine its bid value that maximizes gain $b_k(\pi_k) - p_{jk}$ such that if SU k colludes with other SUs, it fails to minimize the gain. The π_k policy is the decision policy which PUs define for the allocation of white spaces to a truthful SU k in the presence of SUs as bidders. It has been shown that truthful SUs receive higher gain and higher occurrence of winning bids for channel access compared to collusive SUs; while the PUs monetary gain decreases.

2.2.6 Performance enhancement of spectrum leasing schemes

Table 2-2 presents the performance enhancement achieved by the spectrum leasing schemes compared to conventional and traditional approaches in CRNs. The performance metrics are as follows:

P1 *Lower outage probability*. Lower outage probability indicates lesser interruptions of packet transmissions in which transmission does not take place for a certain period of time. For instance, the interruption may be caused by transmission power which is less than a certain threshold value [73], as well as lack of white spaces [74]. Lower outage probability has been shown to enhance QoS (P3) [64].

P2 *Higher outage capacity*. Outage capacity is the maximum achievable transmission rate during the instance outage is occurred. Higher outage capacity indicates higher achievable transmission rate in the presence of outages from time to time, and so it also indicates lower occurrences of outages [73]. Higher outage capacity has been shown to enhance QoS (P3) [67].

P3 *Better QoS level*. Through spectrum leasing, the PUs and SUs achieve QoS enhancement. For instance, higher throughput indicates higher rate of successful data transmission over a channel, which provides better QoS [39]. Higher throughput may also indicate more white spaces, in terms of time duration, being offered to SUs by PUs at a specified cost [48].

P4 *Higher energy efficiency* indicates lower energy consumption by PUs [37]. This is because the SUs help the PUs to relay their packets due to the low channel quality in PUs' direct transmission to PU destination node [69]. With reduced unsuccessful transmission attempts by PUs, the PUs consume lower transmission power and there are more white spaces available to be leased to SUs for monetary gain (P5).

P5 *Higher monetary gain*, which is the gain exclusive for PUs A(1.1). The PUs receive monetary gain as revenue based on the price of the white spaces being offered to SUs through spectrum leasing [42].

P6 *Balanced tradeoff between cost of white spaces and monetary gain*. Generally speaking, the cost of white spaces paid by the SUs is set by the PUs. Higher cost provides higher monetary gain received by PUs at the expense of SUs. Hence, a balanced tradeoff between the cost of white spaces and monetary gain provides a win-win solution for both PUs and SUs [48].

P7 *Balanced tradeoff between PUs' and SUs' channel access time*. Generally speaking, higher channel access time among the PUs may provide better QoS level (P3) among the PUs at the expense of reduced channel access time among SUs, and vice-versa [67]. Hence, a balanced tradeoff between PUs' and SUs' channel access times provides a win-win solution for both PUs and SUs.

P8 *Better security level*. Through the detection of malicious SUs that access PUs' channels in an illegitimate manner, better security level can be achieved contributing to better QoS level (P3) (e.g. throughput) and monetary gain (P5). For instance, in [47], the SUs report their respective channel access time, which is closely monitored by PUs. Hence, malicious SUs that mislead PUs with incorrect information (e.g. channel access time) in order to compete for channel access can be detected by PUs. Subsequently, the PUs evict the malicious SUs from their channels, and this has been shown to achieve higher throughput for PUs and SUs, as well as an increase in PUs' monetary gain.

P9 *Lower PUs' interference level*. Lower interference level to PUs in the use of white spaces by SUs provides better QoS (P3) to PUs. For instance, in [38], a PUs' interference cap, which is the maximum interference level that PUs can tolerate in the use of white spaces by SUs, is set in order to increase PUs' and SUs' throughput performance.

Table 2-2: Performance enhancements achieved by the spectrum leasing schemes

Challenges	References	Performance Enhancement								
		P1 Lower outage probability	P2 Higher outage capacity	P3 Better QoS level	P4 Higher energy efficiency	P5 Higher monetary gain	P6 Balanced tradeoff between spectrum cost and monetary gain	P7 Balanced tradeoff between PUs' and SUs' channel access time	P8 Better security level	P9 Lower PUs' interference level
H1 Increasing the monetary gain of PUs	(Alptekin and Bener, [48])			×		×	×			
	(Lin and Fang, [57])					×				
	(Yi et al., [42])			×		×				
	(Kim and Shin, [58])					×				
	(Song and Lin, [45])			×					×	
	(Wu et al., [36])					×				×
H2 Selecting an optimal channel with white spaces by SUs	(Chan et al., [49])			×		×	×	×		
	(Vilar et al., [52])			×	×					
	(Cao et al., [39])			×						
	(Jayaweera et al., [37])			×	×			×		
	(Murawski and Ekici, [59])			×						
	(Toroujeni et al., [60])			×					×	
H3 Scheduling the channel access of PUs and SUs	(Chen et al., [61])			×				×		
	(Huang et al., [50])			×			×			
	(Lei et al., [62])				×			×		
	(Stanojev et al., [63])	×		×						
	(Wang et al., [64])			×	×			×		
	(Zhang et al., [65])			×	×			×		
	(Zhu et al., [66])			×						
	(Asaduzzaman et al., [67])	×	×					×		
	(Khalil et al., [68])			×						
	(Zhou et al., [43])							×		
H4 Continuous monitoring of white spaces being leased to SUs by PUs	(Jayaweera et al., [9])	×		×					×	
	(Jayaweera and Li, [51])			×					×	
	(Hakim et al., [69])	×		×	×					
	(Sodagari et al., [47])			×		×				×

2.3 Open issues in SL for CRN

This section discusses important open issues that can be pursued in this research area.

2.3.1 Enhancing auction and coordination mechanisms

Generally speaking, auction enhances the performance matrices (i.e. better QoS level (P3) and higher monetary gain among PUs (P5)), and it requires proper coordination in which the PUs (or SUs) make decisions on the selection of SUs (or PUs) participating in spectrum leasing so that both PUs and SUs mutually agree to fulfill each other requirements. For instance, in [37], the PUs choose the SUs that allocates higher transmission power to relay PUs' packets based on the bid values received from SUs through auction. The disadvantages are that, the PUs incur high energy consumption while exchanging control messages and making decisions on the outcomes of auctions. Hence, a third-party auctioneer has been proposed to receive control messages from both PUs and SUs, as well as to make decisions on the auction outcomes [47]. Additionally, the purpose-built third-party auctioneer may reduce latency associated with auction because of the auction being its main and only task. Further investigation can be pursued to investigate a balanced tradeoff between energy consumption and monetary gain in order to enhance the network performance of both the networks in the presence of a third-party auctioneer.

2.3.2 Investigating distributed spectrum leasing schemes

Current research focuses on centralized networks (C1.1) in which PU BS and SU BS exist; however, this may not be the case in distributed networks (C1.2), and so further investigation can be pursued to investigate spectrum leasing in distributed networks. While there are investigations into distributed SU networks [37], this is not the case for PU networks in which most schemes in the literature assume the presence of a PU BS or a single PU node pair. The major challenge in distributed SU networks is that SU BS does not exist, and so the SUs must coordinate among themselves to determine a control channel for the purpose of control message exchange in spectrum leasing. The control channel is important for the exchange of control messages for spectrum leasing. The lack of a control channel has been investigated based on the assumption that the SUs are equipped with learning capabilities [37]; specifically through past experience. Further investigation can be pursued to relax this assumption.

2.3.3 Implementation of security measures

Generally speaking, the implementation of security measures to prevent malicious SUs by PUs may increase the performance matrices (e.g. better QoS level (P3) and higher monetary gain by PUs (P5)). Since the PUs can provide continuous monitoring on SUs' channel access the PUs can detect malicious SUs. The challenge is to reduce the additional overheads, such as energy consumption, incurred by the PUs. This is particularly important because malicious SUs may access the channel (white spaces) in an illegitimate manner, and this minimizes the amount of white spaces for genuine SUs, which subsequently degrades the performance of PUs and SUs. Three examples of security vulnerabilities associated with spectrum leasing are as follows:

- SUs attempt to acquire the white spaces from PUs in an illegitimate manner through untruthfully raising their respective bid values (e.g. SU's transmission power used to relay PUs' packets) [47].
- The winning SUs may further sublease their channels to losing SUs for monetary gain [47].

- The SUs may launch collusion attacks in which SUs participating in an auction collaboratively reduce their bid values that may significantly reduce the monetary gain (P5) of PUs [36].

Further investigations can be pursued to address the aforementioned security vulnerabilities.

2.3.4 Investigating energy-efficient spectrum leasing schemes

In spectrum leasing, the SUs may serves as relay nodes to transmit both PUs' and SUs' transmission packets; hence, they incur higher energy consumption. However, current literature primarily focuses on reducing energy consumption at PUs [37, 64]; and so further investigation can be pursued to reduce energy consumption at SUs. By reducing the transmission power at SUs, there are two main advantages as follows:

- Firstly, it reduces the interference to PUs and its neighboring SUs, and this helps to enhance the PUs' and SUs' performance (e.g. better QoS level (P3)).
- Secondly, it reduces SUs' monetary cost, which may be related to energy consumption used to relay PUs' packets [64].

Further investigation can be pursued to achieve a balanced tradeoff in order to utilize the channel and energy in an efficient manner.

2.3.5 Investigating common assumptions of spectrum leasing

Future investigation can be pursued to relax the following common assumptions, as well as their effects, applied to the investigation of spectrum leasing in CRNs:

- Each node is equipped with two transceivers, namely control transceiver and data transceiver. The control transceiver is always tuned to a single common control channel, which is available at all times; however, the existence of a common channel among nodes may not be realistic [45].
- Each SU observes the similar white spaces, and the transmission from each SU can be observed by all the other SUs [45]. This assumption may not be realistic because each SU may observe different white spaces.
- Each SU BS makes decision on spectrum leasing. For instance, in [37], the SU BS makes decision for SUs' participation in spectrum leasing. However, the presence of a SU BS as a decision maker may not be feasible in distributed networks. There has been very limited literature on distributed approaches (see subsection 2.3.2).

2.3.6 Defining the selection and eviction criterion of SUs by PUs

Generally speaking, there has been very limited research on the selection and eviction criterion of SUs, which are used by PUs. This helps PUs to enhance the overall QoS performance (P3) of PUs' and SUs' networks. Two types of selection and eviction criterion are as follows:

- PUs may allocate white spaces to SUs that demand higher amount of white spaces in order to maximize their respective throughput and the monetary gain; while neglecting other SUs that demand lower amount of white spaces.
- PUs may monitor the SUs' activities so that PUs can evacuate SUs who breach the spectrum leasing contract upon negotiations [58].

Therefore, further investigation can be pursued to define the selection and eviction criterion in order to achieve higher network performance.

2.3.7 Implementation of hybrid model

Generally speaking, there has been limited research on the enhancement of QoS performance (P3) along with the monetary gain received by PUs (P5) in spectrum leasing. In the current literature, the *exclusive-use* model has been widely used in which PUs shares their white spaces to SUs on lease for a definite period of time but cannot re-claim these white spaces even if the PUs encountered the shortage of spectrum. Whereas, in [58] Kim and Shin (2009) propose a *hybrid* model comprised of a shared-use model and an exclusive-use model. In shared-use model, SUs opportunistically use the spectrum while there is no advantage for PUs, neither in terms of monetary gain nor as an improvement of PU network enhancement. The inclusion of shared-use model gives PUs an additional privilege to evict the SUs whenever the PUs needs the white spaces for their own transmission. The challenge arises in hybrid model is the suspension of white spaces to SUs that on one end crucial for the PUs in order to fulfills their spectrum shortage while on the other hand it deteriorates the SU packet transmission which decreases the PU monetary gain. Further investigation can be pursued to investigate a balanced tradeoff that fulfills the PUs spectrum shortage as well as to ensure the minimum transmission requirements of SUs.

2.4 An overview of CR-based testbed implementations

Majority of the research related to CR networks has been limited to theoretical framework [18, 19], and simulation studies [23-25]. In recent years, some essential CR functions, such as channel sensing, have been implemented on real testbeds focusing on PHY and MAC layers [75-79]. However, there is only perfunctory effort to investigate the network layer through real testbed implementations mainly due to the limitations, which discussed briefly in chapter 5. For instance, one of the limitation is only a few nodes have been utilized in network layer implementation [26-28].

In the network layer of CR networks, there are two widely adopted network architectures, namely distributed and centralized models. In the distributed model, each SU node has local spectrum knowledge, which represents the local availability of various channels over time and space [80]. In the centralized model, there is a centralized entity, such as a SU base station. The SU base station has full spectrum knowledge in the form of spectrum occupancy map, which represents the network-wide availability of various channels over time and space. The SU base station sends updates on the spectrum occupancy map to the rest of the SUs [81]. In this work, both centralized and distributed models are considered. In the centralized model, a SL approach is adopted in which the PUs share their spectrum occupancy map with SUs [41, 82]. In the distributed model, PUs do not share the spectrum occupancy map with SUs, and so the SUs must sense for available channels and infer their available time.

Tremendous work has been done to investigate route selection in centralized and distributed models in CR networks using simulation tools (e.g., Qualnet and NS2) [16], [83-85]. However, there has only limited work on route selection conducted on real testbeds (e.g., USRP/ GNU radio) [26-27]; and hence, this is the focus of this thesis. Various route selection approaches have been proposed with the objectives of maximizing throughput [13], [86], minimizing end-to-end delay [87-88], maximizing route stability [27], [89], and minimizing route recovery/ maintenance cost [90].

Three main spectrum-aware route selection schemes that adopt the distributed model have been implemented on a CR testbed. Firstly, in [28], Nagaraju et al. (2010) implemented a joint cross-layer routing and channel selection scheme on a testbed comprised of three USRP SU nodes to select a next-hop node in a single-hop CR network. Generally speaking, in a single-hop network, a single source node selects one of the two destination nodes based on throughput performance. There are two possible routes in the network. The route selection scheme uses the signal-to-interference-noise ratio, which is dependent on binary phase shift keying and quadrature phase shift keying modulation schemes, to achieve the objectives of maximizing throughput and minimizing the interference with neighboring PUs and SUs. Secondly, in [27], Huang et al. (2011) implemented Coolest Path on a testbed comprising six USRP SU nodes to find a stable route traversing across a multi-hop CR network. There are three possible routes in the network. Coolest Path uses channel availability as the routing metric, which is dependent on the number of channel and route switches, as well as the number of route breakages, in order to achieve the objectives of maximizing throughput and minimizing route recovery. Lastly, in [26], Sun et al. (2014) investigated various route selection schemes, particularly SAMER and CRP, on a testbed comprised of up to six USRP SU nodes to find a stable route traversing across a multi-hop CR network. There are four possible routes in the network. The mechanism uses the number of channel switches and the number of route breakages as the routing metrics, which are dependent on channel availability time, in order to achieve the objectives of maximizing throughput and minimizing route recovery.

2.5 Chapter summary

This chapter presents a comprehensive review on spectrum leasing schemes along with the advantages, functionalities, characteristics and challenges of each scheme in CR networks. Spectrum leasing schemes have been shown to address the concerns poised to the traditional CR networks, so that PUs can enhance their network performance and maximize their monetary gain; while the SUs can enhance their network performance through exclusive access to white spaces. Examples of PU's gains are monetary gain and network performance enhancement; while example of SU's gain is dedicated channel access. To achieve these gains, PUs need to determine the cost of the white spaces, the PU's and SU's channel access time, SU's selection as a relay nodes, as well as PU's own packet transmission; while SUs need to select the appropriate PUs according to the SUs' QoS requirements and the cost of white spaces, as well as to determine channel access time between SUs. In the literature, the network topology of PUs and SUs can be either centralized or distributed; and the PUs and SUs operate among themselves using intra-cooperative and inter-cooperative modes, respectively. The challenges associated with PUs are the selection of the appropriate SUs to increase the monetary gain, the distribution of channel access time between PUs and SUs as well as continuous monitoring of SUs' activities; while the challenge associated to SUs is the selection of optimal channels in order to reap the benefits of spectrum leasing. Additionally, various performance enhancement achieved by the spectrum leasing schemes (e.g. lower outage probability and higher outage capacity) are discussed. Further, some open issues are recommend in order to spark new interests in this research area (e.g. enhancing auction and coordination mechanism and investigation of energy-efficient spectrum leasing schemes), as well as new kinds of CR networks such as CR sensor networks. Finally, the end of this chapter discusses the limited work implemented on the real testbed environment for multi-hop CRN; specifically at the network layer.

With the help of literature review of spectrum leasing schemes it has been observed that the research has been done either through designing the mathematical model or by using the simulation tools. However, to the best of my knowledge no work has been initiated earlier to investigate the concept of spectrum leasing on a testbed environment, although some work has been done to investigate the implementation of CR schemes covering the aspects of PHY and MAC layers, as well as the network layer using a testbed environment. So, the review of spectrum leasing schemes revealing the importance of this thesis work, as it serves as a pioneer work that implement the concept of SL scheme on a real testbed environment; which can be shown in the below table 2-3.

Table 2-3: Summary and novelty of this thesis

Work	Physical or MAC layers	Network Layer	Network Size	Multi-hop CRN	Leveraging AI concept	Spectrum Leasing
Nagaraju et al. [28]		×	3	No	No	No
Huang et al. [27]		×	6	Yes	No	No
Sun et al. [26]		×	6	Yes	No	No
Bogale and Vandendorpe [91]	×		2	No	No	No
Zhao et al. [92]	×		2	No	No	No
Reyes et al. [93]	×		2	No	No	No
Qi et al. [94]	×		3	No	No	No
Reddy et al. [95]	×		3	No	No	No
This Thesis	**×**	**×**	**Up to 10**	**Yes**	**Yes, RL**	**Yes**

Next, the research framework and methodology is discussed in the later chapters (i.e., chapter 3, chapter 4, and chapter 5), which can be categorized under one of the following categories:

i. Preliminary evaluation

ii. Multi-hop CRN: An Experimental evaluation.

The earlier category (i.e., preliminary evaluation) discusses a simple SL approach in chapter 3, in which PUs select SUs as relay nodes, on a real testbed environment, as well as using a simulation tool. The PU source node has to select a reliable SU node as a relay node to transfer the packets to PU destination node. The PUs and SUs nodes are configured both on a real testbed environment, as well as using a Matlab tool. Whereas, the latter category (i.e., Multi-hop CRN: An Experimental evaluation) discusses the real testbed implementation for channel selection in chapter 4 and route selection in chapter 5, respectively. The channel and route selection implementations have been implemented in the presence of multi-hop nodes using USRP/ GNU Radio. Note that, the primarily focused of this thesis is to investigate the concept of CRN and SL on a real testbed environment with the incorporation of RL as an artificial intelligence technique.

3.0 PRELIMINARY STUDIES

This chapter presents the preliminary evaluation work that examine the concept of SL in which the PU's needs to select SUs as relay nodes using the simulation tool (e.g., Matlab); as well as deployed it on a real testbed (e.g., USRP/ GNU radio). The functionalities are modeled and solved using Reinforcement Learning (RL). As defined earlier, the primary focused of this thesis is to investigate the concept of CRN and SL on a real testbed environment with the incorporation of RL as an artificial intelligence technique. However, at preliminary evaluation simulation results are also computed alongside of real testbed results for verification purpose (i.e., to see the trend of simulation and testbed results), whereas the later chapters discusses the results related to the real testbed implementation. This chapter helps to design and build an environment to examine the concept of SL without the involvement of multi-hop communication in the network using simulation and experimental evaluations. This provides the foundation to extend this work to more complex multi-hop CR networks on a real testbed environment. This chapter is divided into three sections. Section 3.1 presents an overview of RL technique. Section 3.2 presents the selection criterion of SUs as relay nodes by PUs; this section discusses setup and results that have been simulated and experimented using Matlab and USRP/ GNU Radio implementation to demonstrate the concept of spectrum leasing in CR network. Finally, section 3.3 presents the chapter summary.

3.1 Reinforcement learning

This section presents an overview of one of the main approach used in machine learning i.e. reinforcement learning. In artificial intelligence, machine learning is used as an important tool that allows the machine (or system) to act without any prior or explicit knowledge by performing several tasks; such as recognition, diagnosis, planning and prediction to learn the operating environment of machines. The machine learning can be classified into many different approaches, but in this thesis RL is a selective approach of machine learning.

Reinforcement learning, a pioneered algorithm used in machine learning. RL is initially derived from biologically-inspired learning paradigm but later widely used in artificial intelligence. RL is an intelligent and unsupervised approach because it interacts with the system environment without any prior knowledge of the system. In RL, an agent makes a decision to select an optimal action through trial-and-error learning model [96] from its operating environment in order to achieve an optimal policy that maximizes the agent reward [97]. The major aim of RL is to produce cumulative rewards after the selection of an action. The challenge that experience by an agent is that it may exploit its previous knowledge but also need to explore new action and state spaces from its operating environment. This challenge of an agent is also known as exploitation versus exploration [98].

Generally speaking, RL model is used to solve Markov Decision Processes (MDP) problems, which usually have following elements [99], shown in Figure 3.1.

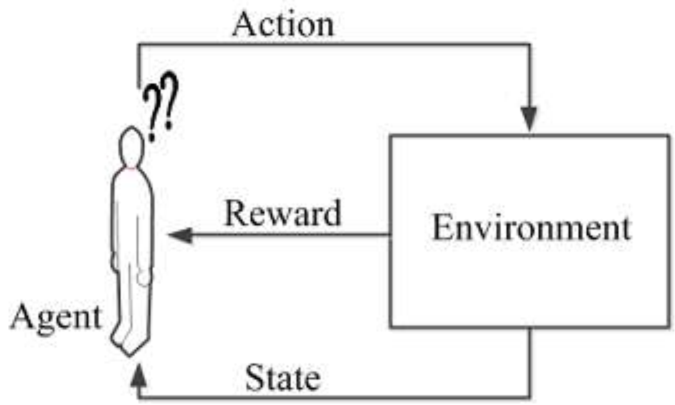


Figure 3-1: General RL model

- State, S is the possible condition of the operating environment.
- Action, A is the set of permissible actions.
- Rewards, $R : S \times A$ is the performance enhancement, which is a function of state and action spaces, received by the agent.

In MDP, the state and action spaces are normally infinite but can be used with finite number of state and action spaces [100]. The RL equation is computed using Q-learning technique as discussed in [55]. The Q-learning technique is dependent on Q-value, which is a function of state s_t^i and action a_t^i in order to compute the delayed rewards $r_{t+1}^i(s_{t+1}^i) \in R$. The Q-value $Q_t^i(s_t^i, a_t^i)$ can be updated as time goes by using Q-function as shown in Equation (3.1).

$$Q_{t+1}^i(s_t^i, a_t^i) \leftarrow (1-\alpha)Q_t^i(s_t^i, a_t^i) + \alpha[r_{t+1}^i(s_{t+1}^i) + \gamma \max_{a \in A} Q_t^i\left(s_{t+1}^i, a\right)] \qquad (3.1)$$

where, $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq 1$ are the learning rate and discount factor, respectively. Higher learning rate α indicates higher speed of learning from the operating environment;

whereas, higher discount factor γ indicates greater dependency on the long-term future rewards $\gamma \max_{a \in A} Q_t^i\left(s_{t+1}^i, a\right)$ compared to delayed reward $r_{t+1}^i\left(s_{t+1}^i\right)$.

3.2 Relay node selection for spectrum leasing in CR network

This section presents the scenario that demonstrate the concept of spectrum leasing in CR network. In this scenario, PUs want to select SUs in order to achieve the following:

- The licensed users (or primary users, PUs) and unlicensed users (or secondary users, SUs) interact with each other to achieve mutual agreement on channel access in order to increase their respective network performance.
- The PUs must select suitable SUs as relay nodes which are expected to uphold the leasing agreement.

Generally speaking, the SU's transmission power must fulfill the minimum and maximum power threshold levels imposed by PUs. The minimum power thresholds ensure that a satisfactory level of successful transmission can be achieved by SUs while helping to relay PUs' packets. On the other hand, the maximum power threshold ensures that SUs' interference to PUs is acceptable to PUs. In this section, the PUs announce their requirements on minimum and maximum power threshold levels to SUs for the selection of relay nodes; while the SUs maintain their respective transmission power within the threshold level defined by PUs in order to increase their respective network performance (e.g. throughput and end-to-end delay performances). The functionalities are modeled and solved using Reinforcement Learning (RL), which determines the suitable SUs as relay nodes on the basis of the aforementioned power threshold criterion. Previously in the literature, many researchers investigate the effects of interference and the ways to mitigate it in order to enhance the performance of PUs' and SUs' using RL in traditional CR networks. For instance, in [101-103], the authors propose RL framework to evaluate the effects of interference generated by SUs on the performance of PUs' and the ways to mitigate it. Whereas, a very limited research have been carried out to evaluate and mitigate the effects of interference on spectrum leasing using RL framework. For instance, in [37], the authors proposed RL framework for the auction of PUs spectrum to SUs in a cooperative manner, in order to exploit the spectrum leasing that primarily aim is to provide the power efficiency for PU network, which in a broader aspect reduces the effect of SUs interference on PU network. On the other hand game theory framework has mostly been used to evaluate and mitigate the effects of interference on spectrum leasing both in centralized and distributed network environments. For instance, in [104], the authors proposed the game-theoretic framework to evaluate the effects of interference within the SUs in Ad Hoc networks. Furthermore, in [9], [51] and [105], the authors advocates the use of game-theoretic framework for PUs and SUs, in which PUs adjust their tolerable interference level that SUs need to fulfilled. In this work, the PUs announce their requirement on minimum and maximum power threshold levels to SUs for the selection of relay nodes; while the SUs maintain their respective transmission power within the threshold level defined by PUs in order to increase their respective network performance. The functionalities are modeled and solved using RL (see section 3.1), which determines the suitable SUs as relay nodes on the basis of the aforementioned power threshold criterion.

The preliminary results show that the number of SUs that qualify as relay nodes increases with the maximum power level imposed by PU, and thus it is expected to provide PUs' and SUs' performance enhancement (e.g. throughput). It also shows that, the convergence rate of SUs' power level increases with the number of simulation iterations. The section is divided in 2 subsections. The first subsection presents a problem and the proposed system model; whereas, the second subsection presents the discussions on the simulation and experimental results.

3.2.1 Problem and proposed system model

In spectrum leasing, a PU wants to lease its spectrum to the appropriate SUs that strictly fulfills the conditions of agreement in order to improve network efficacy. However, study [58] revealed that SUs may breach the agreement of spectrum leasing and attempt to acquire extra network resources from PUs in an illegitimate manner. In this scenario, PUs can drop those SUs in order to safeguard their own interest as well as to serve the rightful

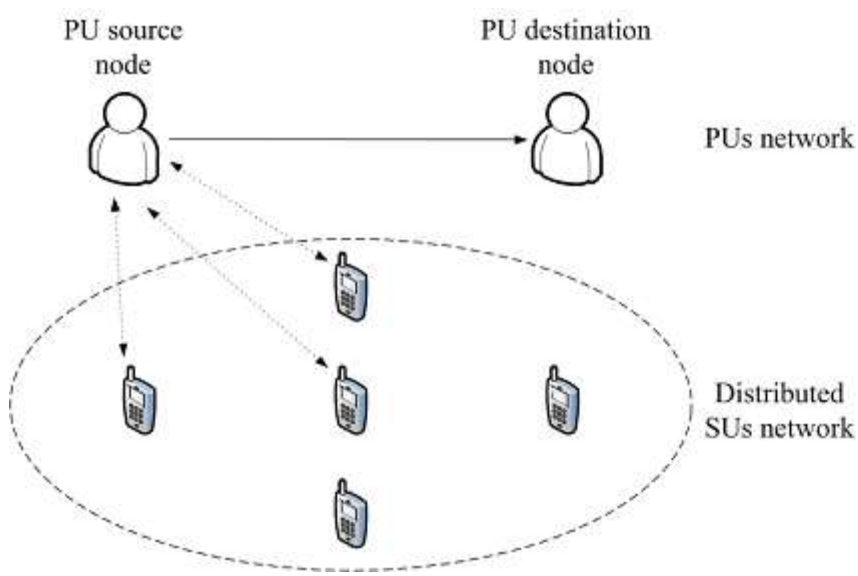


Figure 3-2: Proposed system model

SUs in the CR networks. Hence, it is critical to develop a model that encourages SUs to follow the spectrum leasing agreement, and upon breaching of the agreement, PUs can evict them from the CR networks.

The proposed system model, consider a scenario of spectrum leasing in which a PU source and destination node pair and there are k SUs as shown in Figure 3.2. Initially PU source node defines the maximum and minimum levels of interference that it can tolerate from SUs. In return, SUs submit their bids to PU source node, in which the bids indicate their respective transmission power for transmitting the PUs' and the SUs' packets. The PU source node analyzes the bids, and selects SUs, which transmit the PUs' and SUs' packets using the transmission power which falls in between the defined maximum and minimum threshold power levels, as relay nodes.

The proposed RL model shown in Table 3-1, aims to select the SUs as relay nodes, which fulfill the minimum and maximum power threshold levels imposed by PUs. The state s_t^i represents a SUs' power level. The action a_t^i represents the selection of a SU node k, which offers a bid value within the range of maximum and minimum power threshold values. The reward of a state-action pair r_t^i (s_t^i, a_t^i) represents the cost incurred in collaboration with PU.

Table 3-1: RL model at PU source node

State	State $s_t^i = \left(P_{1.t}^i, P_{2.t}^i, \ldots, P_{K.t}^i\right) \in S$, where $P_{k.t}^i \subseteq \{P_1^i, P_2^i, \ldots, P_K^i\}$ represents the bidding value received from the neighboring SU $k \in K$ which indicates their respective power level, and K represents all SU nodes located within the transmission range of PU source node
Action	Action $a_t^i \in A = \{1,2, \ldots, K\}$ represents the selection of SU neighbor node k, which offers a bid value within the range of maximum and minimum power threshold values to collaborate
Reward	Reward $r_{t+1}^i\left(s_t^i, a_t^i\right) \in \{-1,1\}$ represents a constant value to be rewarded to all SU nodes. Value 1 indicates that a PU node i does not select a SU $a_t^i = k$ as relay node due to unfulfilled power threshold criterion set by the PU, while value -1 indicates that a PU node i selects a SU $a_t^i = k$ as relay node

3.2.2 Results and discussions

This section, investigate the proposed scheme described in subsection 3.2.1, with the help of Matlab and USRP/ GNU Radio for simulation and implementation purpose, respectively. The number of PUs and SUs are 2 and 10, respectively. The PU source node must define the maximum and minimum power levels as threshold values, which SUs must fulfil by adjusting their respective transmission powers within the threshold values.

The collaboration and selection of SUs by PU source node is formulated using RL as discussed in subsection 3.2.1. The SU relay selection is dependent on the bid values of SUs', which represent their respective transmission power. If the transmission power levels of SUs are within the threshold values defined by PU source node, the PU source node selects these SUs as relay nodes for their own packet transmission. The simulation parameters are shown in Table 3-2.

Table 3-2: Simulation parameters

Parameters	**Value**
Number of PUs	2
Number of SUs	10
PU power threshold level 1	Upper level: 0.3
	Lower level: 0.1
PU power threshold level 2	Upper level: 0.6
	Lower level: 0.4
RL parameters	Learning rate (α): 0.2
	Discount factor (γ): 0
No. of simulation iterations (10 times)	v_1 (PU power threshold level 1)
	v_3 (PU power threshold level 2)
No. of simulation iterations (30 times)	v_2 (PU power threshold level 1)
	v_4 (PU power threshold level 2)

The PU source node sets these maximum and minimum power levels for SUs as the percentage of its own maximum transmission power. For instance, Figure 3.3 defines different values of SU maximum power levels, which are basically the fraction of the PU maximum power level (i.e. 0.2, 0.4, 0.6, 0.8 and 0.9). The PU maximum power level is

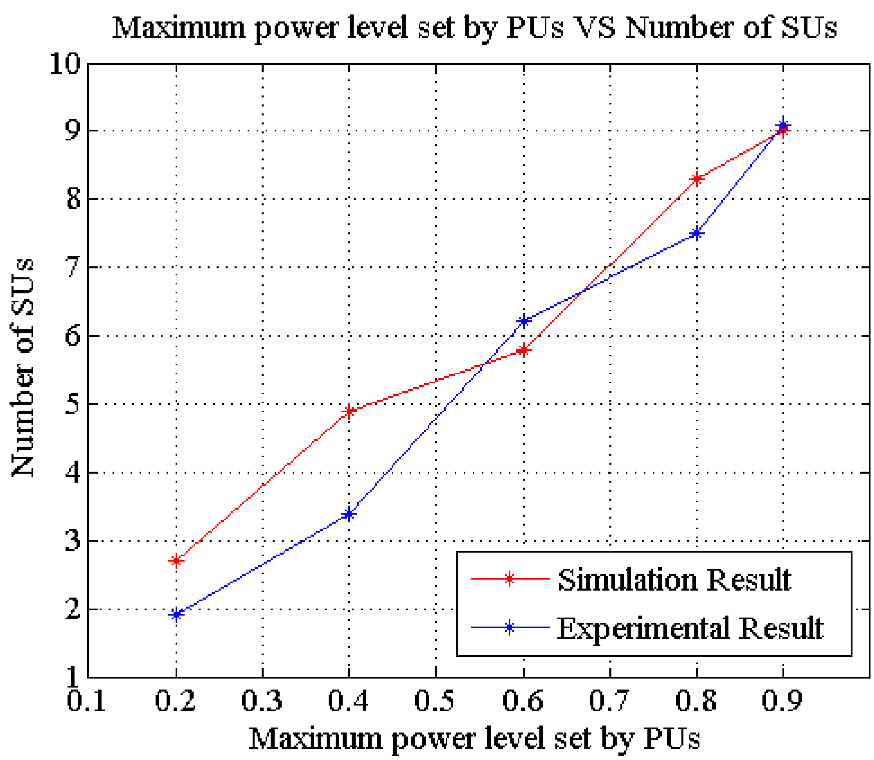


Figure 3-3: Impact of increasing the maximum power levels on the numbers of SUs

assumed to be 1. These aforementioned values are set to examine the impact of increasing the maximum power level on the numbers of SUs which can be shown in the form of percentage in Table 3-3. It has been shown that when PU source node increases its maximum power level for SUs, the number of SUs also increases significantly. For instance, the simulation result (see the red line) reveals that when the maximum power level is at 0.2 of the PU power, the system can accommodate 28% of SUs; while the percentage of SUs is significantly increases to 49% when the maximum power level is set at 0.4 of the PU power. Similarly, the experimental result (see the blue line) reveals that when the maximum power level is at 0.2 of the PU power, the system can accommodate 19% of SUs; while the percentage of SUs is significantly increases to 34% when the maximum power level is set at 0.4 of the PU power. In summary, RL has been shown to select the suitable SUs according to the PU requirements.

Table 3-3: Impact of increasing the maximum power level on the numbers of SUs

		Maximum power level set by PUs				
		0.2	0.4	0.6	0.8	0.9
Number of SUs accommodate in spectrum leasing (system model) w.r.t the total number of SUs (in percentage)	Simulation (Matlab)	28%	49%	58%	82%	90%
	Testbed (USRP/ GNU Radio)	19%	34%	62%	75%	91%

Figure 3.4 and Figure 3.5 show that simulation result and experimental result taken at two different power threshold values. The first threshold value upper level is 0.3 and lower level is 0.1; while the second threshold value upper level is 0.6 and lower level is 0.4, respectively. In the preliminary evaluation, the maximum upper (power) level set by the PU for SUs is 0.6, which indicates that PU is more interested to use SUs as its relay for the transmission of its own packets. Furthermore, values v_1 and v_3 represent the average of 10 simulation iterations; whereas values v_2 and v_4 represent the average of 30 simulation iterations. It can be shown that as the number of simulation iteration increases, the successful SUs powers are become more stable. This has been with the aid of RL, which selects suitable SUs with the predefined maximum and minimum threshold values set by PUs.

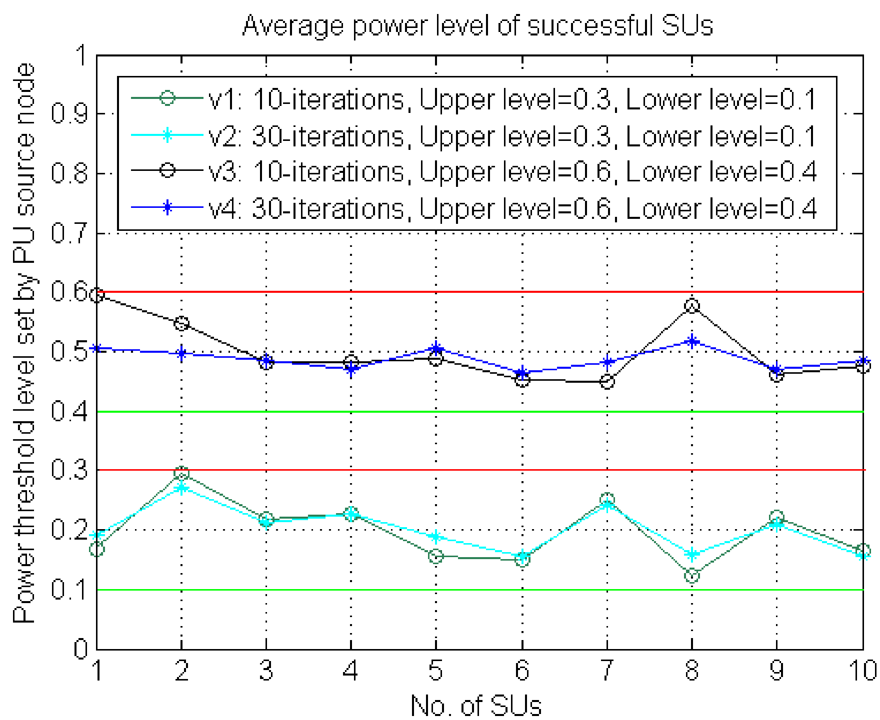


Figure 3-4: Power level of SUs after 10 and 30 iterations at two different power threshold values defined by PUs (simulation results)

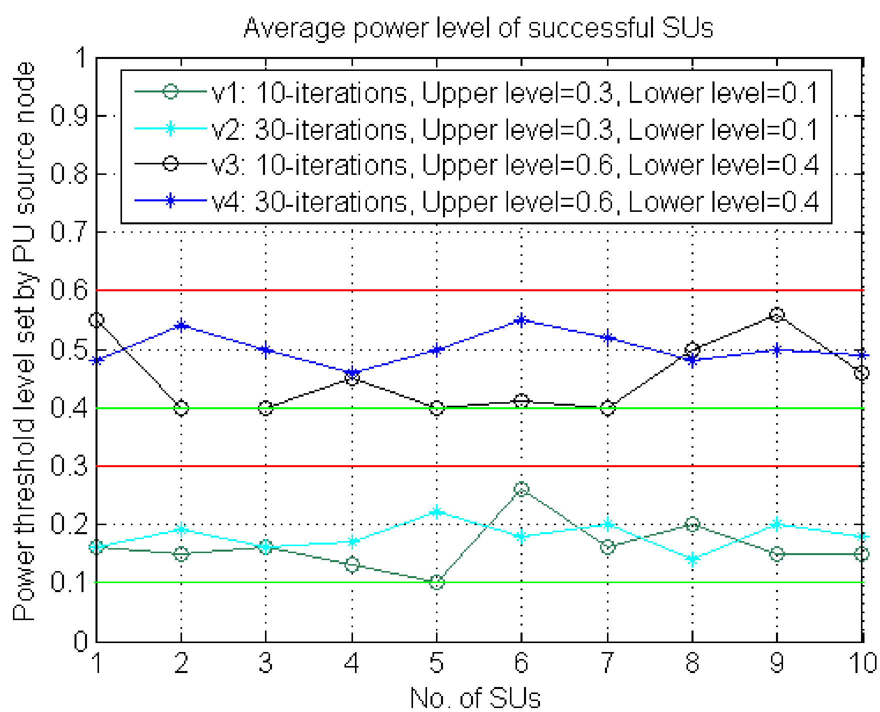


Figure 3-5: Power level of SUs after 10 and 30 iterations at two different power threshold values defined by PUs (experimental results)

3.3 Chapter summary

In this chapter, a distributed secondary user network is proposed that can utilize the allocated spectrum of primary user network in a mutually collaborative manner by exploiting the concept of spectrum leasing. Furthermore, reinforcement learning is incorporated into the system model, which aims to select the suitable SUs as relay nodes for spectrum leasing in cognitive radio networks. The results show that, the number of successful SUs that fulfill the requirement imposed by PUs increases with the increase in the maximum threshold power level defined by PUs. Additionally, transmission power of successful SUs stabilizes as time goes by.

4.0 CHANNEL SELECTION IN MULTI-HOP CRN: AN EXPERIMENTAL EVALUATION

The next two chapters present the work that has been implemented on a real testbed environment for channel selection and route selection in a multi-hop CRN, respectively; whereas no simulation work has been simulated in these chapters, as the scope of this thesis is on the experimental implementation. A video file is used to evaluate the performance of the aforementioned schemes (i.e., channel selection and route selection). It is due to the reason that CR can be used to protect and monitor the video surveillance application, which has been used in many places (e.g. offices, schools, banks, military sites, roads etc.) primarily for safety, protection and monitoring purposes. It has been reported in the literature that the video traffic is expected to increase by a factor of 65 times, and thus become the dominant source of data traffic [106]. Furthermore, many researchers have used video traffic to analyze the performance of their proposed ideas in CR network [107-110]. In this chapter, the deployment of channel selection scheme has been discussed that provides a proof-of-concept for the selection of channels by SUs in a multi-hop CRN using the experimental testbed (e.g., USRP/ GNU radio). This work provides the foundation to extend the work to more complex multi-hop CR networks in the next chapter. Reinforcement learning (RL), which is an artificial intelligence approach, is applied to select and switch to the best possible operating channel. This chapter is divided into three sections. Section 4.1 presents the introduction. Section 4.2 presents the experimental setup and results using USRP/ GNU radio in a multi-hop CR environment, in which SU nodes sense the environment and select a channel that offers least PUs activity. Finally, section 4.3 presents the chapter summary.

4.1 Introduction

In the CR network context, the unlicensed users (or secondary users, SUs) want to use the licensed users' (or primary users', PUs') channels in order to secure maximum transmission time for packet transmission. In this investigation, the multi-hop network comprises of three SUs, namely source, relay and destination nodes. There is a single PU. In recent years the core concepts of CR, namely spectrum sensing has been implemented using a testbed environment. For instance, in [111-116], the authors demonstrated one of the core concepts of CR, namely spectrum sensing, which is used to detect the available spectrum holes for SUs transmission. In [117], the authors implemented the DSA algorithm for the transmission of video using USRP and experimental result shows that DSA improves the performance of a video application in comparison with traditional environment (i.e. non-DSA). In [118], the authors implemented the CR system for WiMAX and LTE; which are emerging wireless technologies. In [119], the authors implemented the CR system over frequency modulation (FM) bands. This section, implement a RL-based channel selection mechanism on a multi-hop CR network that enables SUs to evade PUs' activities. The implementation results show that RL is an effective approach that helps SUs to select the optimal channel, and so improves their network performance.

4.2 Channel selection: setup and results

The section is divided in 2 subsections. The first subsection presents an experimental setup; whereas, the second subsections presents the discussions on the experimental results.

4.2.1 Experimental setup

The experimental setup comprises of two computers, four USRP/GNU units, and a gigabit Ethernet switch as shown in Figure 4.1. All USRP units are connected through a switch to one of the computers. The switch serves as a convergence point for the computers and the USRPs, so it is not necessary to connect each USRP unit to a computer. Three SUs and a single PU are configured through USRP which serve as the hardware platform. Each USRP has two transceivers, and so it can transmit and receive packets simultaneously. The two computers, namely the source computer and the destination computer, are installed with Ubuntu Linux operating system and GNU Radio, which serve as the software platform for USRPs.

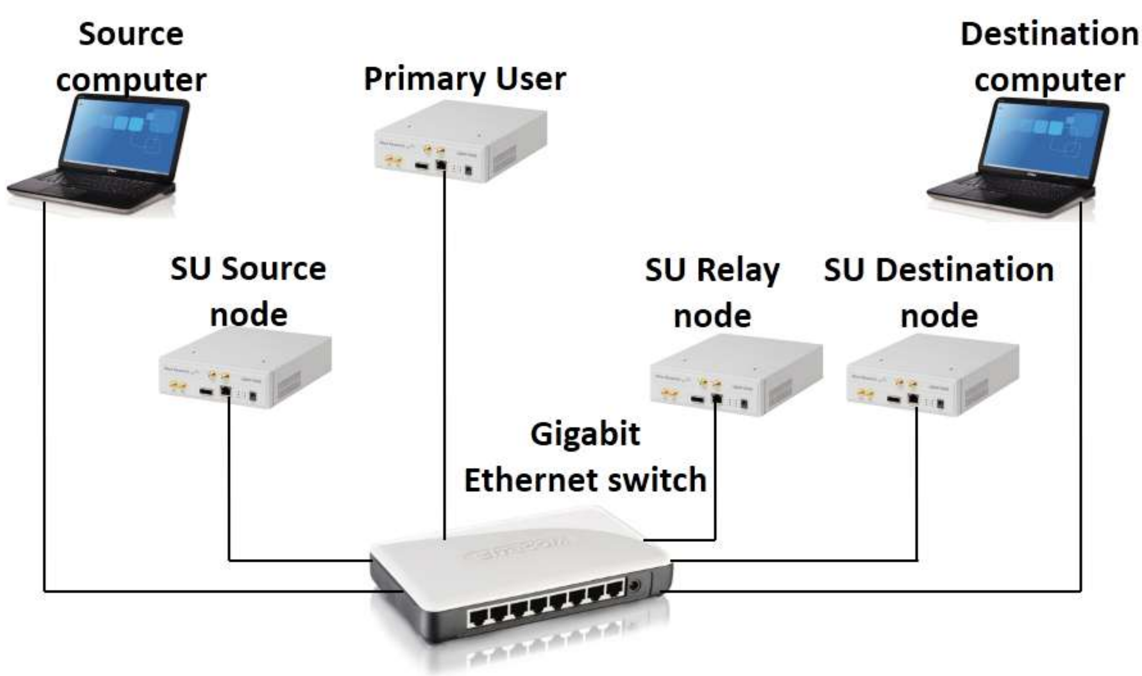


Figure 4-1: System model

The SUs form a multi-hop network, and they are the source, relay and destination nodes, and transmit video packets from the source computer to the destination computer. The transmission from the source computer to the destination computer is accomplished through two wireless channels, specifically one channel for source to relay transmission and the other for relay to destination transmission, respectively. This is to avoid contention among the SUs as each node can transmit and receive packets at the same time. The PU transmits in any of the available channels.

A 20 MHz-wide spectrum is divided into five different channels, namely channels 1 to 5. Either channel 1 or 2 is used for transmissions from the source node to the relay node; and either channel 3, 4 or 5 is used for transmissions from the relay node to the destination node. This allocation of reserved channels to the respective nodes allows them to select the next operating channel whenever the current operating channel is used by the PUs. This means that the source to relay transmission may use either channel 1 or 2, and this depends on the channel selected by a PU. However, the relay to destination transmission is not straightforward because there are three available channels, and the PU tends to appear more often in one channel than the other. In this case, RL is used for the selection of the next operating channel, which is less likely to be used by the PUs. The experimental parameters are shown in Table 4-1.

Table 4-1: Experimental parameters

Parameters	Value
Number of SUs	3
Number of PUs	1
Number of channels	5
Learning rate, α	$\{0.2,\ 0.8\}$
Discount factor, γ	0

Selection of the operating channel for the transmission of video packets from the relay node to the destination node is based on the RL approach. Figure 4.2 shows the RL model embedded in the SU relay and SU destination nodes. Generally speaking, each of them senses the three available channels, and each channel receives either a reward or a punishment, which is solely dependent on the presence PU activities on the respective channel. For instance, if the current operating channel for the relay to destination transmission is channel 3, and PU activities re-appear on this channel, then this channel gets punishment, and channels 4 and 5 receive reward. Based on the Q-values of the channels, a channel with the highest Q-value is selected as the next operating channel. The new RL equation, which is derived from Equation (3.1), can be expressed as:

$$Q_{i,t+1} = (1-\alpha)Q_{i,t} + \alpha r \qquad (4.1)$$

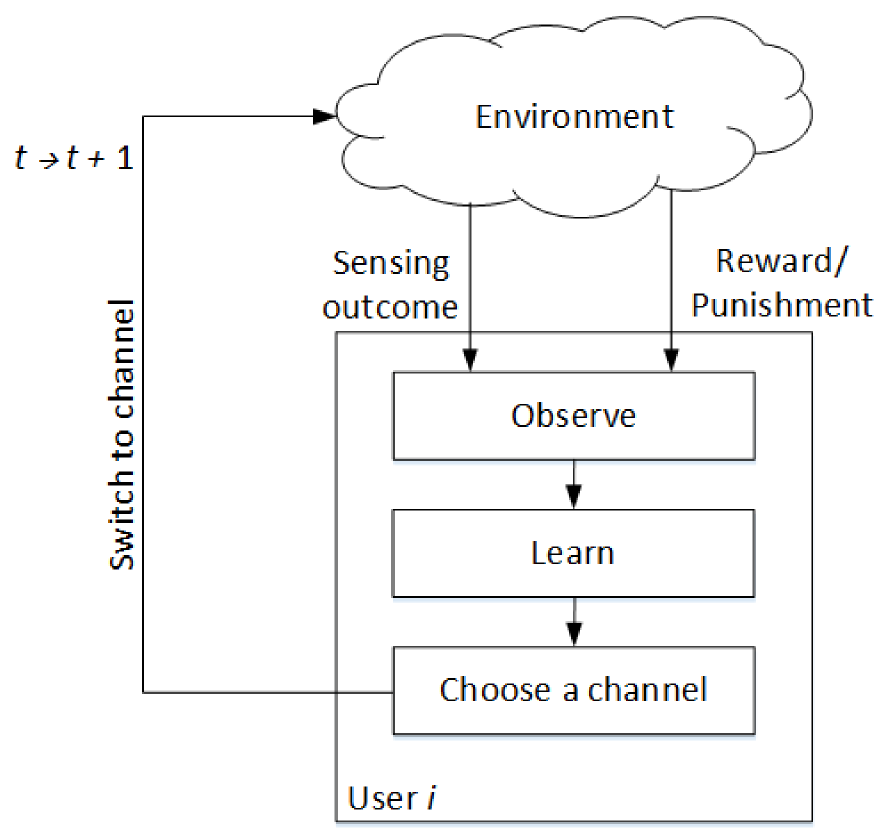


Figure 4-2: RL model for channel selection

In Equation (4.1), the discount factor $\gamma = 0$, and other parameters are shown in Table 4-1. Equation (4.1) updates the Q-value of each channel (i.e. channels 3, 4 and 5) after every time window of 1 minute; which is further decomposed into transmission and channel sensing time periods, as shown in Figure 4.3. The channel switch and selection criterion is solely dependent on the PUs' activities. The transmission signals are detected through energy detection. The energy detection technique [84] follows Equation (4.2).

$$x_{i,t} = \begin{cases} n_i(t), H_0 \\ h_i(t).\, e_i(t) + n_i(t), H_1 \end{cases} \tag{4.2}$$

where $n_i(t)$ is the noise signal, and $h_i(t).\, e_i(t)$ is the transmission signals. H_0 represents the absence of a signal, while H_1 represents the presence of a signal.

4.2.2 Experimental results

This work implements RL-based channel selection mechanism on a multi-hop CR network in order to switch to the best possible operating channel with the objective of evading the PUs' activities. Firstly, it shows the effects of different learning rate α of the RL approach on the learning performance. Secondly, it compares the throughput performance achieved by both RL and non-RL approaches.

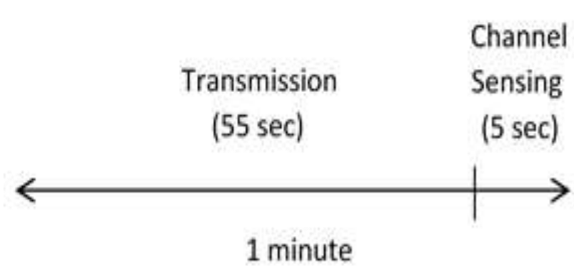


Figure 4-3: Time slot allocation

Figures 4.4 and 4.5 show the Q-values (or accumulated reward value) of the operating channels 3, 4 and 5, which are used in the relay to destination transmission, as time goes by. Two different learning rates $\alpha = 0.2$ and $\alpha = 0.8$ are shown in Figures 4.4 and 4.5, respectively. Note that, as shown in Equation (4.1), higher learning rate indicates more dependence on the previous Q-value and less responsive to the delayed reward. For instance, in Figure 4.4 and Figure 4.5, at time instant 10, channel 5 is selected for transmission as it has the highest accumulated reward value among the available channels. Furthermore, it can be observed that in Figure 4.4, when the learning rate is $\alpha = 0.2$, the learning mechanism achieves faster convergence than that in Figure 3.9, where the learning rate is $\alpha = 0.8$. However, in Figure 4.4, the accumulated reward value has large fluctuations and less stable compared to that in Figure 4.5. Hence, while lower learning rate increases convergence rate, the fluctuation can be high, and so it is advisable to use higher learning rate after the system has achieved convergence.

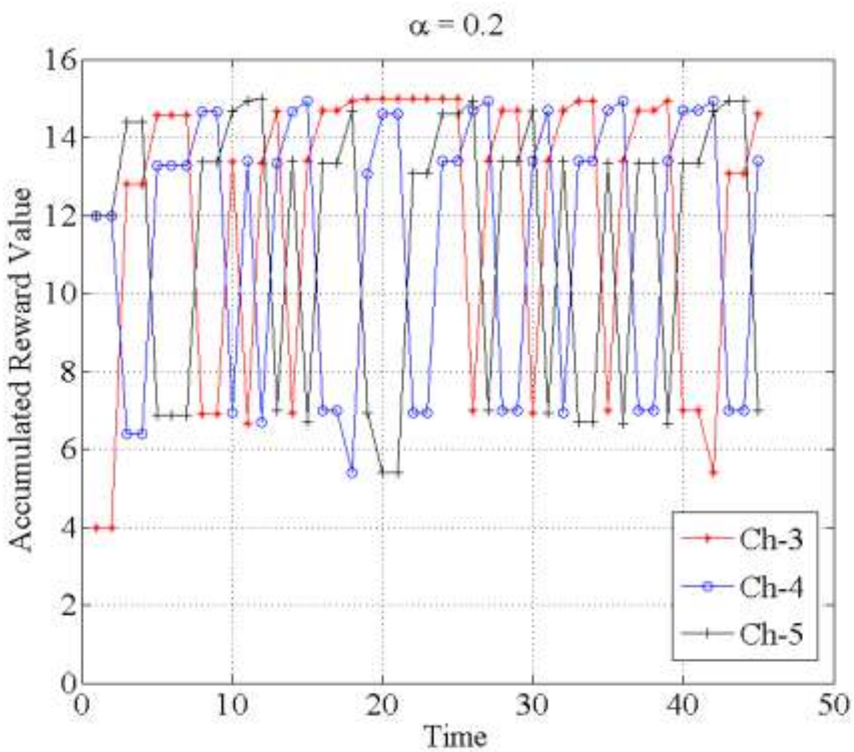


Figure 4-4: Accumulated reward with α = 0.2

Figures 4.6 and 4.7 show the throughput performance with respect to time for RL and non-RL approaches. Using the RL approach, whenever the transmission channel among the SUs is re-occupied with signals from PUs, the SUs update the Q-value of the available channels, and switch to another channel free from the PUs' activities. Subsequently, the SUs resume transmissions in the newly chosen operating channel. Without using the RL approach, fixed channels (i.e. channel 2 and channel 3) are permanently allocated for the transmissions among the SUs. In Figure 4.6, 50% of the transmissions from PUs occur in the fixed channels; while in Figure 4.7, 67% of the transmissions from PUs occur in the fixed channels. It can be seen that, in Figures 4.6 and 4.7, whenever the channel is re-occupied, transmission is interrupted and comes to a halt. Using RL, the throughput performance is stable with respect to time, and it is comparatively higher in both Figures 4.6 and 4.7.

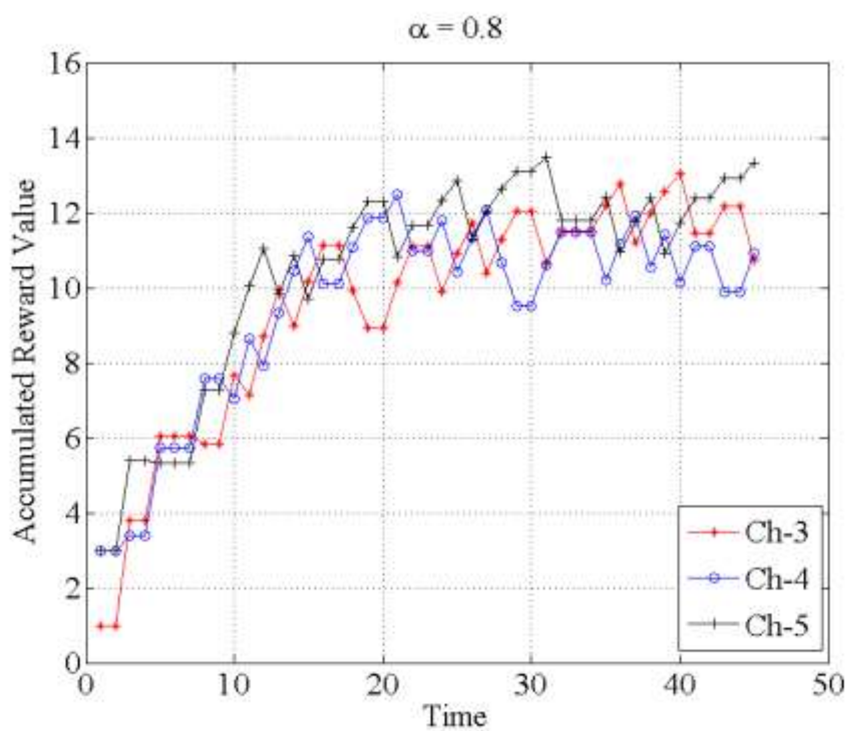


Figure 4-5: Accumulated reward with α = 0.8

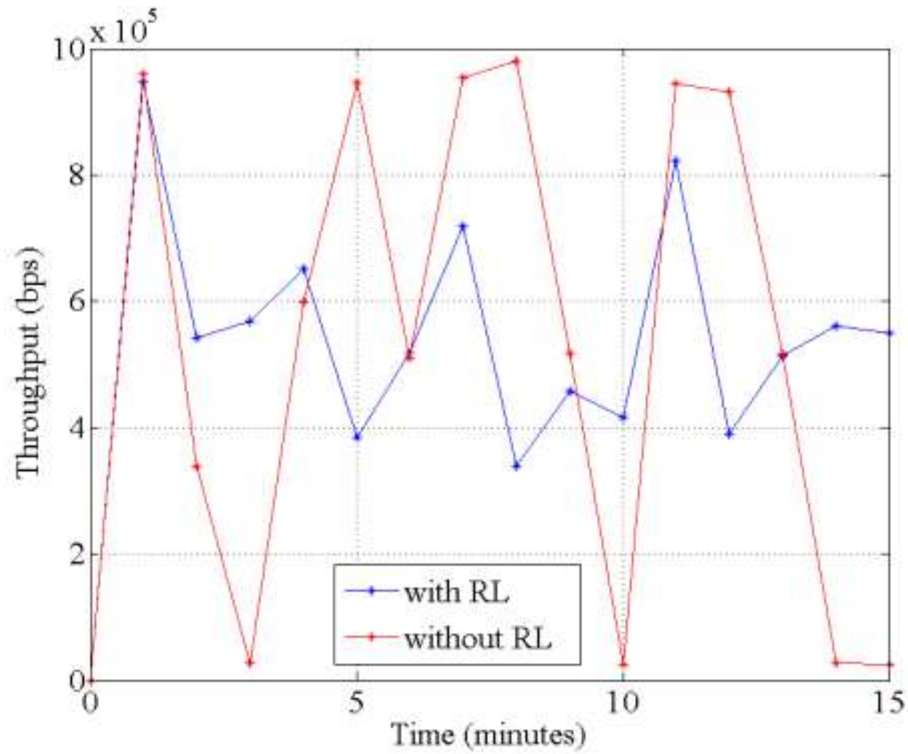


Figure 4-6: Throughput performance when PUs activities on the channels is 50%

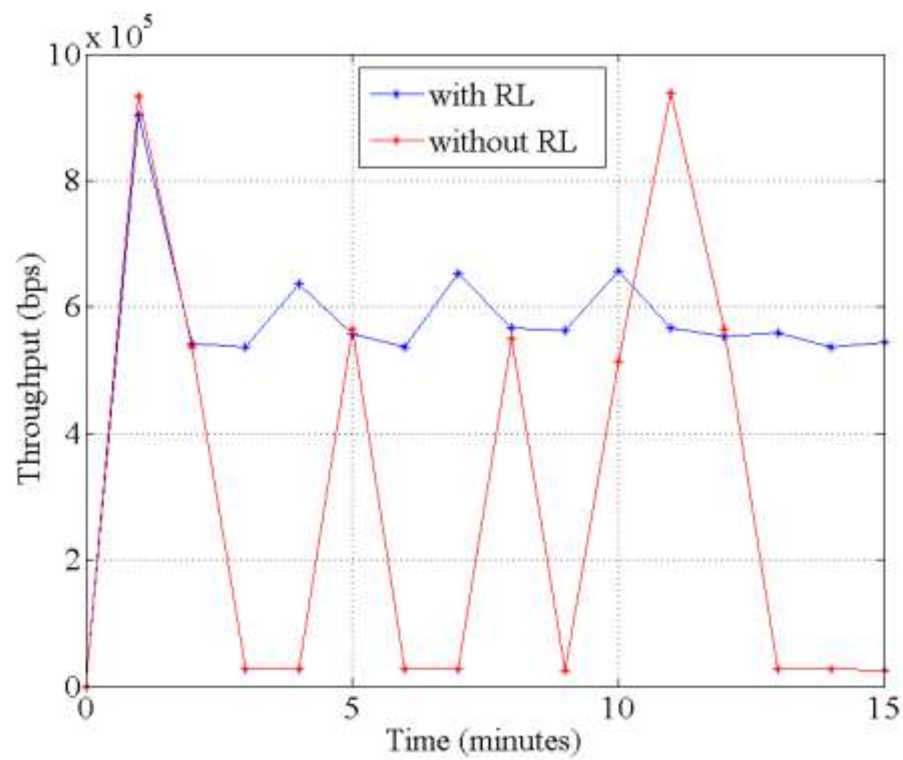


Figure 4-7: Throughput performance when PUs activities on the channels is 67%

4.3 Chapter summary

In this chapter, a reinforcement learning (RL)-based channel selection mechanism is implemented on a multi-hop cognitive radio (CR) network in order to select the best possible operating channel. Experimental results show that smaller learning rate α increases convergence rate but the accumulated reward value is unstable. Also, throughput performance is stable in the RL approach compared to that in the non-RL approach. This is because SUs switch to the next operating channel free from PUs' activities. This scheme is suitable to be applied in a video surveillance system to evade denial of service attack (DoS) launched by the malicious users.

5.0 ROUTE SELECTION IN MULTI-HOP CRN: AN EXPERIMENTAL EVALUATION

This chapter presents the work that has been implemented on a real testbed environment for the selection of route in a multi-hop CRN. The multi-hop CRN has drawn significant research interest in recent years, although majority of the work has been validated through simulation. A key challenge in multi-hop CR network is to select a route with high quality of service (QoS) and lesser number of route breakages. This chapter proposes three route selection schemes to enhance the network performance of CR networks, and investigate them using a real testbed environment, which consists of universal software radio peripheral (USRP) and GNU radio units. Two schemes are based on reinforcement learning (RL), while a scheme is based on spectrum leasing (SL). RL is an artificial intelligence technique, whereas SL is a new paradigm that allows communication between licensed and unlicensed users in CR networks. The proposed route selection schemes are compared with a route selection scheme in the literature, called highest-channel (HC), in a multi-hop CR network. With respect to the QoS parameters (i.e., throughput, packet delivery ratio and the number of route breakages), the experimental results show that RL approaches achieve better performance in comparison with the HC approach, and also achieve close to the performance achieved by the SL approach. This chapter is divided into five sections. Section 5.1 presents an introduction. Section 5.2 presents contribution. Section 5.3 presents system architecture. Section 5.4 presents route selection. Section 5.5 presents experiment and evaluation. Finally, section 5.6 presents the chapter summary.

5.1 Introduction

As discussed, majority of the research related to CR networks has been limited to theoretical framework [18-19], and simulation studies [23-25]. In recent years, some essential CR functions, such as channel sensing, have been implemented on real testbeds focusing on PHY and MAC layers [75-79]. However, there is only perfunctory effort to investigate the network layer through real testbed implementations, and there are three main limitations. Firstly, only a few nodes have been utilized in existing network-layer implementations [26-28]. Secondly, monetary constraint has discouraged network-layer implementation as more nodes and computing resources are needed to investigate multi-hop transmission. Thirdly, the underlying layers (i.e., physical and data link layers), particularly the hardware and software processing delays, can affect the network layer performance [26, 27], [120, 121]. This means that the choice of hardware and software for the underlying physical and data link layers can significantly affect the network-layer performance, which is undesirable. This work addresses the limitations associated with the network-layer implementation using a simplified system architecture to construct a larger network consisting up to ten nodes in which the universal software radio peripheral (USRP) hosts are directly connected to a single computer using an Ethernet switch. This allows the extension of existing implementations [26, 27] by increasing the number of USRP hosts in the platform. The Ethernet switch reduces the effects of latency so that the performance of network-layer schemes can be analyzed.

In this chapter, an experimental setup has been deployed to examine route selection schemes in multi-hop CR platform using USRP [114], [122], and GNU radio toolkit [123]. Generally speaking, USRP, which is an off-the-shelf wireless host, enables each SU to autonomously and dynamically configure various operating parameters, such as channel frequency and modulation scheme, for data transmission using a GNU radio software application. GNU radio, which is an open source software platform, generates signals for USRP nodes and performs waveform-specific processes including modulation (e.g., GMSK), as well as packet encoding and decoding. Three route selection schemes are deployed based on: 1) the traditional reinforcement learning (RL) approach, 2) a RL approach with average Q-value, and 3) a spectrum leasing (SL) approach. RL is an artificial intelligence technique in which a decision maker (or an agent) learns about its operating environment and makes decisions on action selection that provides system performance enhancement without using prior or explicit knowledge. SL is a new paradigm that allows communication among PUs and SUs in CR networks. The proposed schemes select the best possible route from a SU source node to a SU destination node in a multi-hop CR network in order to improve QoS parameters (i.e., throughput and packet delivery ratio) and routing stability (e.g., the number of route breakages).

5.2 Chapter contribution

The main contribution of this chapter is to propose and implement three route selection schemes based on RL and SL on real testbed environment using USRP/ GNU radio platform with the objective of improving QoS performance in multi-hop CR networks (see section 1.5, point 3), while taking into consideration the limitations of the underlying USRP/ GNU radio platform for network-layer implementation. To the best of my knowledge, this is the first testbed implementation of RL-based and SL-based route selection schemes in CR networks, taking into consideration the limitations of network-layer implementation.

A summary of the notations used in this chapter is shown in Table 5-1.

Table 5-1: Summary of notation

Category	Notations	Description
Network	$P = \{1,2,\dots,\lvert P\rvert\}$	A set of PUs
	$M = \{m_1, \dots, m_N\}$	A set of USRP SU nodes
	m_1	SU source node
	$X_{h,j_h \in J_h} = \{m_2, \dots, m_{N-1}\}$	A set of SU intermediate nodes, where $h = \{1,2,\dots,\lvert H\rvert\}$ is the number of hops from SU source node m_1, and $j_h \in J_h$ is one of the h-hop nodes from the SU source node m_1
	$m_{n_{h,j_h}}$	SU intermediate node with node identification number n_{h,j_h}
	m_N	SU destination node
	$K = \{k_1, k_2, \dots, k_k, \dots, k_{\lvert K\rvert}$	A set of routes in the network
	$\mathbb{R}_{m_1,m_N}$	Route record list in the RREQ message sent from a SU source node m_1 to SU destination node m_N
Channel	t_s	Channel sensing time window
	$C = \{c_1, c_2, \dots, c_c, \dots, c_{\lvert C\rvert}\}$	A set of channels in the network
	$\tau^p_{c_c,ON}$	ON duration of PU $p \in P$ in channel $c_c \in C$
	$\tau^p_{c_c,OFF}$	OFF duration PU $p \in P$ in channel $c_c \in C$
	$\lambda^p_{c_c,ON}$	Average ON time of PU p on its channel c_c
	$\lambda^p_{c_c,OFF}$	Average OFF time of PU p on its channel c_c
	L_{k_k}	Set of links in a route k_k
	$\varphi^{i,j,k_k}_{t,c_c,OFF}$	Estimated average channel available time of channel c_c on link between SU node i and SU node j of route k_k at time t for RL-based scheme
	$\mathbb{e}^{i,j,k_k}_{t,c_c,OFF}$	Exact channel available time of channel c_c on link between SU node i and SU node j of route k_k at time t for SL-based scheme
	$\Gamma^{k_k}_{\beta,t}$	Channel available time of bottleneck link in route k_k at time t
RL	$s_t^{m_N}$	The state represents a potential SU destination node at time t
	$a_t^{m_{n_{1,j_1}}}$	The action represents a neighbor node $m_{n_{1,j_1}}$, which is a single hop away from a source node m_1
	$Q_t^{m_1}(s_t^{m_N}, a_t^{1,j_1})$	The Q-value for a state-action pair $(s_t^{m_N}, a_t^{1,j_1})$ calculated at source node m_1 for route k_k at time t
	α	Learning rate, and its range is $0 \le \alpha \le 1$
	a_t^*	The selected optimal action at time t
	$\bar{Q}_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$	The average Q-value is a mean ratio of the sum of all Q-values up to time t

To achieve chapter contribution, three route selection schemes based on RL and SL are proposed and implemented for multi-hop CR networks. The RL and SL approaches address the dynamicity of the PUs' activities (or channel availability) and select the best possible route in a multi-hop CR networks in order to improve QoS parameters, particularly throughput and packet delivery ratio, as well as the number of route breakages, which represents the route stability. Using RL, SUs learn about the average channel available time and select a route that maximizes the SUs' network performance. On the other hand, SL allows PUs to communicate with SUs and lease their channels to them. Hence, the SL approach may be more suitable in a centralized network in which the SUs has direct communication with PUs. In general, SL offers two main advantages. Firstly, it improves the SUs' channel utilization and network performance based on the spectrum occupancy map sent by the PUs to SUs. Secondly, it offers remuneration to PUs in terms of monetary gain or performance enhancement (e.g., SUs help PUs to relay packets [81]). Hence, the main difference between RL and SL is that, SUs are not informed of the channel utilization of PUs in RL, and the SUs are informed of such information in SL.

This chapter implements the RL-based and SL-based route selection schemes in a real testbed environment using USRP/ GNU radio platform. This chapter also proposes a system architecture to address three main limitations associated with network-layer implementation. Firstly, the system architecture can establish a larger network, which is necessary in multi-hop network-layer implementation for a meaningful investigation in contrast to two nodes (i.e., a transmitter and a receiver) in physical-layer implementation and single-hop transmission (i.e., point-to-point and point-to-multipoint) in data link-layer implementation. While route selection mechanism has been investigated in [28], the implementation focuses on single-hop transmission. Secondly, the system architecture addresses the monetary constraint. The requirement to purchase more equipment (i.e., nodes and computing resources) has discouraged researchers to setup a real testbed environment to investigate the network layer. As an example, in [26], each of the six USRP SU nodes must be connected to a single computer. The system architecture uses a switch to connect the USRPs to a single computer (addressing the second limitation), and more USRPs can be connected to the switch to establish a larger network (addressing the first limitation). Thirdly, the system architecture reduces the effects of hardware and software processing delays in USRP/ GNU radio to network-layer performance. If such delays are taken into account, the choice of hardware and software for the underlying physical and data link layers can significantly affect the network-layer performance. This is particularly significant in multi-hop communication as connecting each USRP SU node to a different computer and running separate software code in each computer incur hardware and software delays at each SU intermediate node. In this work the system architecture uses a single computer running a single software code to coordinate the USRPs. As most processes are performed by a single computer, there are no hardware and software processing delays at each SU intermediate node. Furthermore, a switch provides a seamless control message transmission which is out-of-bound in nature, so that the control message transmission is not affected by data transmission. This is important as each USRP module is only equipped with a single transceiver, and so it cannot transmit data and control messages simultaneously.

5.3 System architecture

This section presents the architecture of the USRP/ GNU radio platform for CR networks. Figure 5.1 shows an architecture with six USRP SU nodes; while another architecture with ten USRP SU nodes is shown in Figure 5.11. In Figure 5.1, the SUs are represented as: a source node (m_1), intermediate nodes (m_2, m_3, m_4, m_5), and a destination node (m_6). This experiment shows and compare the performance of several route selection schemes in selecting the best possible route out of a number of routes. In the literature, investigations have been conducted with two possible single-hop routes (between the source node and the destination node) in a network with three USRP SU nodes [28], as well as three [27] and four [26] possible routes with a maximum of three hops in a network with six USRP SU nodes. In this work, investigations are conducted with four possible routes with a maximum of three hops in a network with six USRP SU nodes in the 6-node topology (see Figure

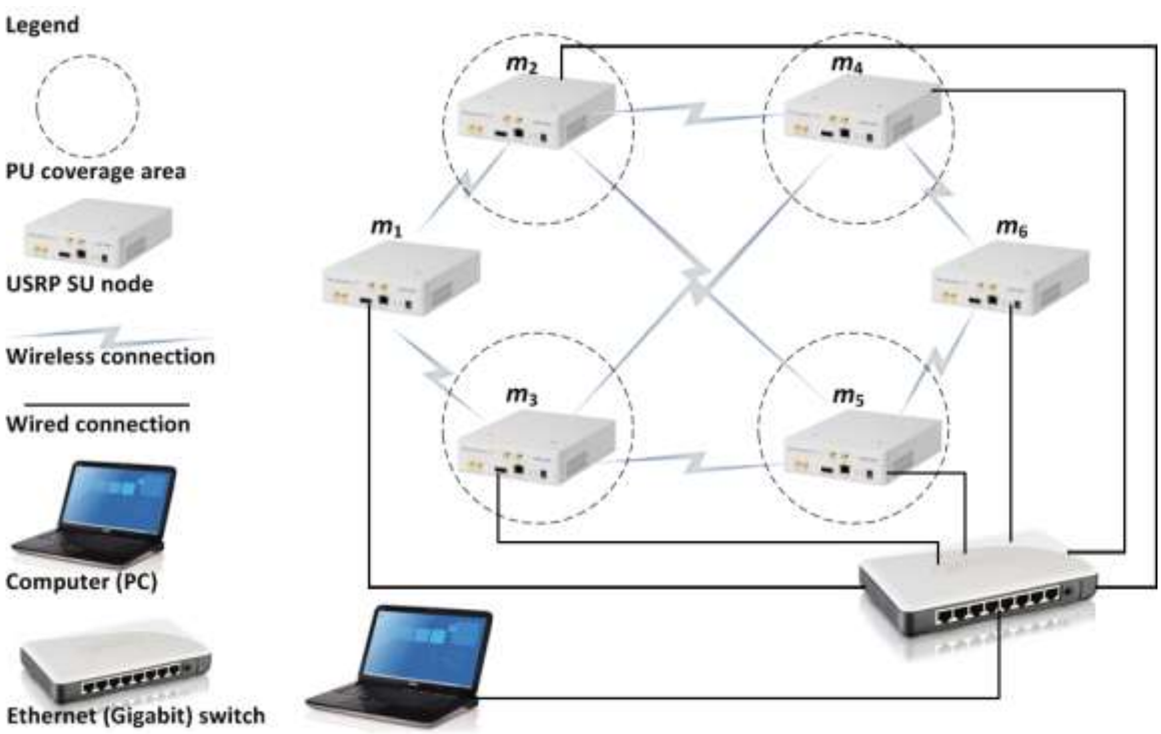


Figure 5-1: A 6-node topology consists of six USRP SU nodes

5.1) and five possible routes with a maximum of four hops in a network with ten USRP SU nodes (see Figure 5.11). The PUs are emulated using a Python script, and their interference is represented by dash-line circles in Figure 5.1 and Figure 5.11. For instance, in Figure 5.1, there are four PUs that interferes with the intermediate nodes (i.e., m_2, m_3, m_4, m_5), and each PU can interfere with three SU links from each of the SUs. Specifically, PU interference at intermediate node m_2 affects links $m_1 - m_2$, $m_2 - m_4$ and $m_2 - m_5$, PU interference at intermediate node m_4 affects links $m_2 - m_4$, $m_3 - m_4$ and $m_4 - m_6$, and so on. The USRP hardware provides the RF frontend. GNU radio is a software interface written in Python, and it is installed in the computer. It serves as a signal processing block that perform tasks such as generating and reconfiguring waveforms. In this implementation, it serves two main purposes. Firstly, it houses the decision making engine, and defines the operating environment (e.g., the PUs' activities, which are exponential ON/OFF processes). Secondly, it provides a software interface to the USRP platform. For instance, it receives the channel state information (e.g., channel number and channel availability) from the operating environment. The decision making engine is one of the main components of GNU radio software, and it obtains decision making factors from the operating environment (e.g., PUs' activities), analyzes the decision making factors, and makes selection of the optimal action (e.g., route selection). Both decision making factors and actions are stored in the knowledge base. More details about the decision making engine for RL-based and SL-based approaches can be found in Sections 5.4.3 and 5.4.4, respectively. The computer and USRP SU nodes are connected to a gigabit Ethernet switch via gigabit wired connections so that the need to connect each USRP SU node to a different computer is not necessary. The gigabit wired connections emulate a common control channel (CCC) used by the USRP SU nodes to exchange route selection messages such as RREQ and RREP. The USRP SU nodes are connected via wireless medium to form a multi-hop CR network. In Figure 5.1, the source node m_1 chooses a route that has higher channel available time to the destination node m_6. Further details of the USRP and GNU Radio are presented in the rest of this section.

5.3.1 USRP unit

Figure 5.2 shows the transmit and receive paths in a USRP unit. There are four main sections. Firstly, the radio frequency (RF) section comprises a set of VERT900 antennas, namely RF1 and RF2, which allows transmission and reception in two different channels, respectively. Each antenna is connected to a WBX transceiver daughterboard. The antenna and transceiver daughterboard can transmit and receive radio signals ranging from 824 MHz to 960 MHz. Secondly, the immediate frequency (IF) section consists of analog/ digital converters, as well as digital up/ down converters. Thirdly, the baseband section performs the proposed route selection schemes. Fourthly, the data section provides a user interface for developing intelligent and knowledge-based mechanisms on the USRP units.

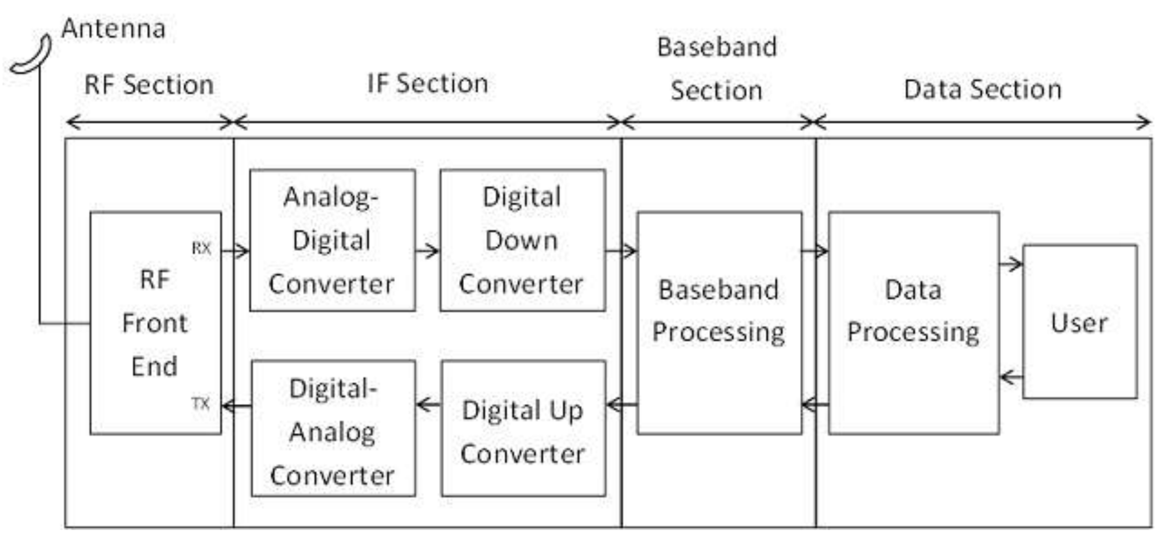
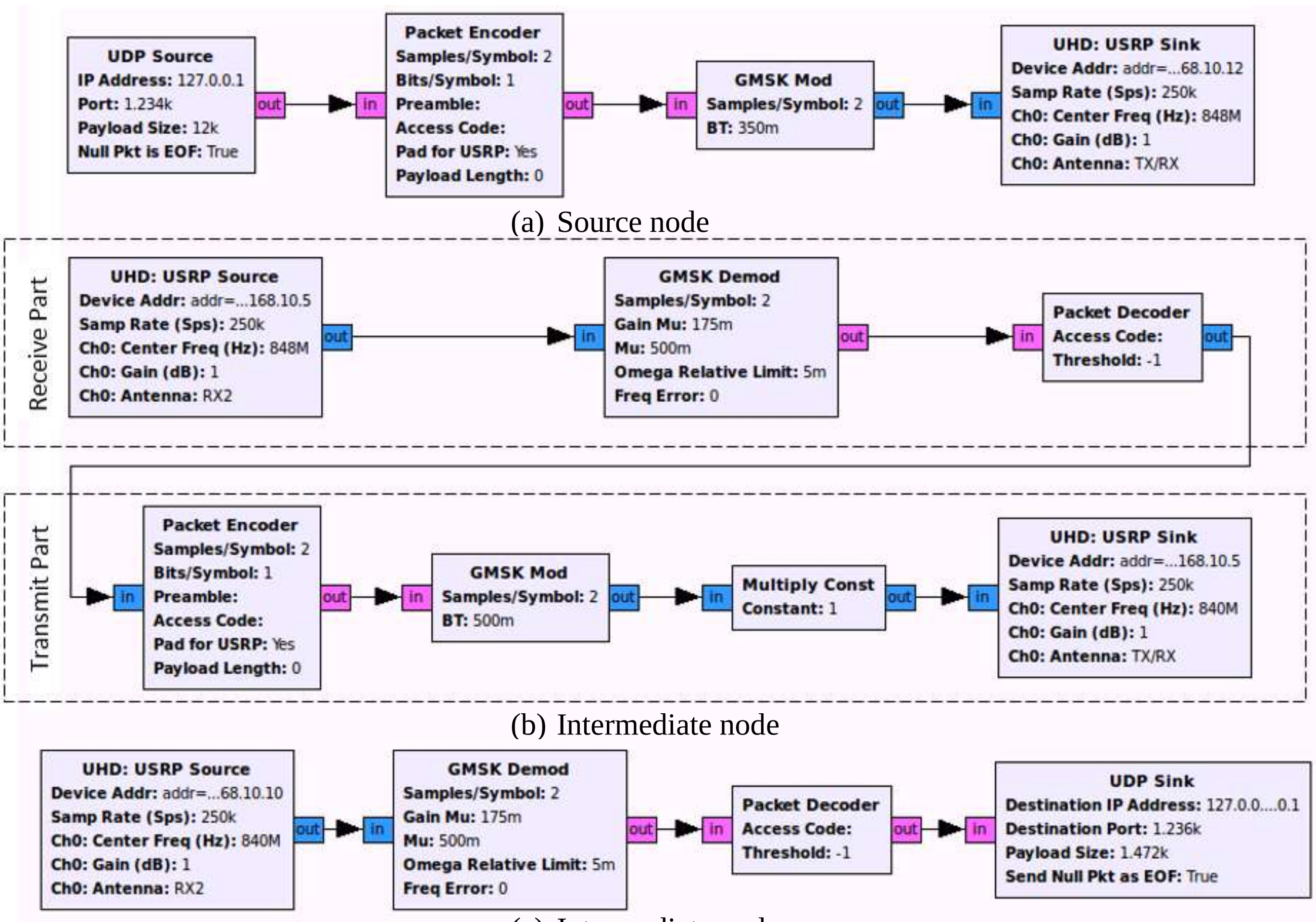


Figure 5-2: Transmit and receive paths in a USRP unit

This enables a SU network to make the right decisions on route selection in order to enhance network performance.

5.3.2 GNU radio

Using GNU radio, the functionalities of the transmitters and receivers are represented as flow graphs. Generally speaking, a flow graph starts with a source (e.g., a user datagram protocol (UDP) source) and ends with a sink (e.g., a USRP sink). The schematic representations of the flow graphs for the source, intermediate, and destination nodes are shown in Figure 5.3(a), 5.3(b) and 5.3(c), respectively. In Figure 5.3(a), the source node initiates packet transmission in which a UDP source receives video frames with a payload size of 12 KB from a computer with an IP address 127.0.0.1 via port 1234. The 'Null Pkt is EOF' is set to 'True' to indicate that the end of file occurs when no packet is received. The UDP source sends the received frames of information, such as video frames, to the packet encoder and modulation blocks. The packet encoder block converts the frames into packets by adding headers, access codes, preamble codes and so on. The 'Samples/Symbol' and 'Bit/Symbol' are set to low values to avoid an error called underrun which occurs when the computer is not fast enough to send video frames to the USRPs. The 'Preamble' and Access Code' values are left empty so that preliminarily data is not needed and access code is not assigned to encoded packets. The GMSK modulation converts the packets into signals. Similarly, 'Samples/Symbol' is set to a low value to avoid underrun. Subsequently, the signals are ready for transmission through the USRP sink block via a TX/RX antenna. Specifically, the USRP SU device, which has an IP address 192.168.10.12, sends the signals at a sampling rate of 250,000 samples/ seconds using a channel frequency of 848 MHz with a channel gain of 1dB. In Figure 5.3(b), the intermediate node receives and



(a) Source node

(b) Intermediate node

(c) Intermediate node

Figure 5-3: Flow graphs for GNU radios

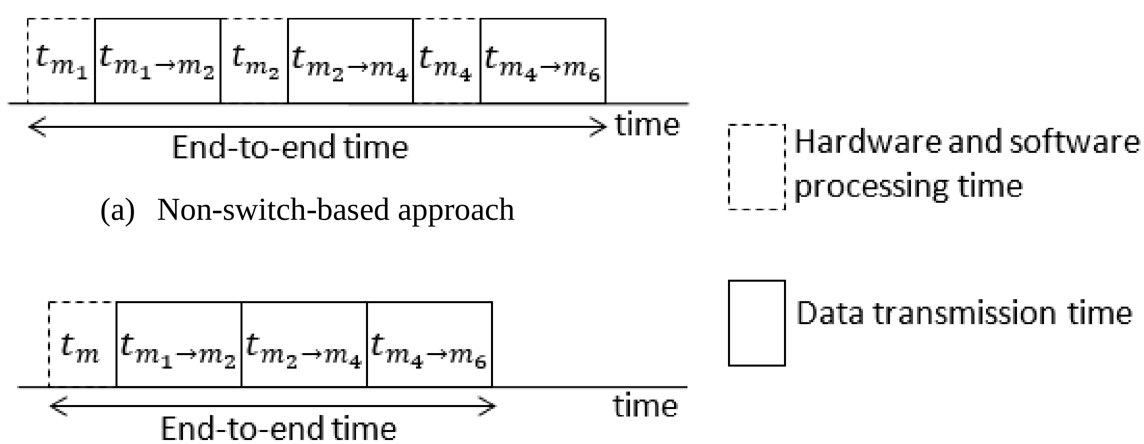
retransmits the signals towards the destination node. The intermediate node has two types of paths, namely receive path and transmit path. In the receive path, the USRP SU node, which has an IP address 192.168.10.5, receives the signals at a sampling rate of 250,000 samples/ seconds via its antenna RX2 using a channel frequency of 848 MHz with a channel gain of 1dB. The signals are then sent to demodulation and packet decoder blocks which convert the signals into packets, and then back into frames. Similarly, ‘Samples/Symbol’ is set to a low value in GMSK Demod to avoid an error called underrun, and ‘Access Code’ value is left empty in Packet Decoder so that access code is not assigned to encoded packets. Next, in the transmit path, packet encoder and modulation blocks reconvert the frames into packets, and modulate them into signals again. Then, the multiply const block amplifies the signals. 'Constant' indicates that the signal power is amplified with the number of times indicated by the value. With a value of 1, there is no amplification as the USRP SU nodes are placed close to each other (see Section 5.5.1). Finally, similar to the USRP sink block in the source node, the USRP sink block in the intermediate node transmits the signals. In Figure 5.3(c), the destination node serves as the sink node. The USRP source block receives the signals and sends them to the demodulation and packet decoder blocks which convert the signals into packets and then back into frames. The UDP sink block captures these frames so that applications, such as a media player and a web browser, receive the information at the destination node.

5.4 Route selection

While routing in multi-hop CR networks has been investigated intensively using simulation platforms (e.g., Qualnet and NS2) [121], there is only perfunctory effort made to implement them on experimental platform. To conduct experiment on route selection in a multi-hop CR network, three schemes are proposed. The first two schemes are based on Q-routing [124], which is a popular approach in RL. Although, it has been widely applied in wireless networks and tested through simulation platforms [124, 125], there is lack of experimental investigation. The third scheme is based on the spectrum leasing concept. One of the major issues of the USRP/ GNU radio platform is the effects of the underlying delay on network performance. The underlying delay is caused by the processing delays of hardware (i.e., reconfigurations and processes of USRPs and the computer) and software (i.e., initialization and processes of GNU radio or Python codes), and it increases with the number of route breakages [121], [126, 127]. This means that higher number of route breakages increases the hardware and software processing time causing a decline in throughput and packet delivery ratio. In [127], the underlying delay is reported as being in the range of 28.9 ms to 36.9 ms. Since the primary focus of this work is on the network layer, and so USRP/ GNU radio platform does not consider the underlying processing delays. Section 5.4.1 fur provides further description about delay in USRP/ GNU radio. Section 5.4.2 presents the system model. Sections 5.4.3 and 5.4.4 present the route selection schemes based on RL and SL, respectively.

5.4.1 An overview of the underlying latency in USRP/ GNU radio platform

Generally speaking, route selection schemes can be implemented on a USRP/ GNU radio testbed using a non-switch-based approach or a switch-based approach. In the non-switch-based approach, each USRP SU node is connected to an individual computer, hence each node incurs the underlying hardware and software delays, which makes it challenging to investigate the network performance achieved by upper layers. In the switch-based



(a) Non-switch-based approach

(b) Switch-based approach

Figure 5-4: Processing time along a route from a source node to a destination node in a multi-hop CR network approach, which is used in this work, USRP SU nodes are connected to a single gigabit Ethernet switch, which is connected to a single computer. Hence, all the nodes along a route use a single Python code in the switch-based approach, instead of their own individual codes in the non-switch-based approach. This helps the switch-based approach to exclude the underlying hardware and software delays found in the non-switch-based approach. Figure 4 shows the difference in the end-to-end time for the two approaches whenever a new route is established. Suppose, the source node m_1 selects a new route $m_1 - m_2 - m_4 - m_6$ to the destination node m_6 in Figure 5.1. Figure 5.4(a) shows the non-switch-based approach in which the hardware and software processing time t_{m_1} is incurred at the source node m_1 for reconfiguration, the data transmission time $t_{m_1 \rightarrow m_2}$ is incurred for data transmission from the source node m_1 to the intermediate node m_2, and so on. Figure 5.4(b) shows the switch-based approach in which the hardware and software processing time t_m is only incurred at the beginning of a route to reconfigure nodes m_1, m_2, m_4 and m_6. Note that, without any changes of route, the hardware and software processing time is not incurred in both non-switch-based and switch-based approaches as reconfiguration is not required. Nevertheless, a route change is necessary due to the reappearance of PUs' activities.

5.4.2 System model

The system model consists a set of PUs $P = \{1,2, \dots, |P|\}$ and a set of available channels $C = \{c_1, c_2, \dots, c_c, \dots, c_{|C|}\}$, where $|P|$ and $c_{|C|}$ represent the number of PUs and channels, respectively. Figure 5.5 shows a network topology, in which the USRP SU nodes are represented by $M = \{m_1, m_2, \dots, m_N\}$. In the network, there is a single USRP SU source node $m_1 \in M$, a set of intermediate nodes $X_{h,j_h} = \{m_2, \dots, m_{N-1}\} \subseteq M$, and a single destination node $m_N \in M$. A set of routes $K = \{k_1, k_2, \dots, k_k, \dots, k_{|K|}\}$ can be established in the network. Each route $k_k \in K$ has a set of links L_k from the source node m_1 to the destination node m_N (e.g., links $m_1 - m_2$, $m_2 - m_4$ and $m_4 - m_6$ in Figure 5.1). In the set of intermediate nodes X_{h,j_h}, an intermediate node $m_{n_{h,j_h}}$ has an identification (ID) $n_{h,j_h \in J_h}$, where $J_h = \{1,2, \dots, |J_h|\}$ is a set of nodes which are h hops from the source node m_1. Specifically, node $m_{n_{h,j_h}}$ is located at $h \in \{1,2, \dots, |H|\}$ hops from the source node m_1, and it is the j_hth node in the set of J_h. The node identification n_{h,j_h} is computed as follows:

$$n_{h,j_h} = h + j_h + \begin{cases} 0 & if\ h = 1 \\ \sum_{h=1}^{h-1} (\ |J_h| - 1) & if\ h > 1 \end{cases} \tag{5.1}$$

The channel selection is dependent on the channel state, which is a two-tuple information comprised of the PU idle/ busy state and the channel available time. The PU activity in each channel is either ON (i.e., busy or PUs' activities appear in the channel) or OFF (i.e., idle or no PUs' activity in the channel) state. The ON duration $\tau^p_{c_c,ON}$ and OFF duration $\tau^p_{c_c,OFF}$ of a PU $p \in P$ in its channel $c_c \in C$ follows a Poisson model, and they are exponentially distributed with rates $\lambda^p_{c_c,ON}$ and $\lambda^p_{c_c,OFF}$, respectively. The terms ON duration and PU-ON time, as well as OFF duration and PU-OFF time, are used interchangeably. This work adopts three assumptions on channel selection and access. Firstly, the underlying channel sensing mechanism of the SUs in the data link layer can sense the channel accurately within a channel sensing time window t_s [128] used by PUs to estimate the channel available time of each PUs' channel in longer term. Note that, the time horizon is segregated into time windows, each of which is segregated into channel sensing time window t_s and data transmission time window t_d. Secondly, the neighboring links of SUs use distinct channels in order to avoid data link-layer interference among the respective SUs [129]. Thirdly, as the effects of typical phenomena like fading and shadowing have been well investigated in the literature [130, 131], in this work the focus is on the main characteristic of CR, which is the dynamicity of PUs' activities, so the USRP SU nodes can be placed close to each other as shown in Figure 5.10 while emulating the CR environment. These assumptions are adopted to simplify the underlying physical and data link layers as the focus of this work is on the network layer and the main characteristics

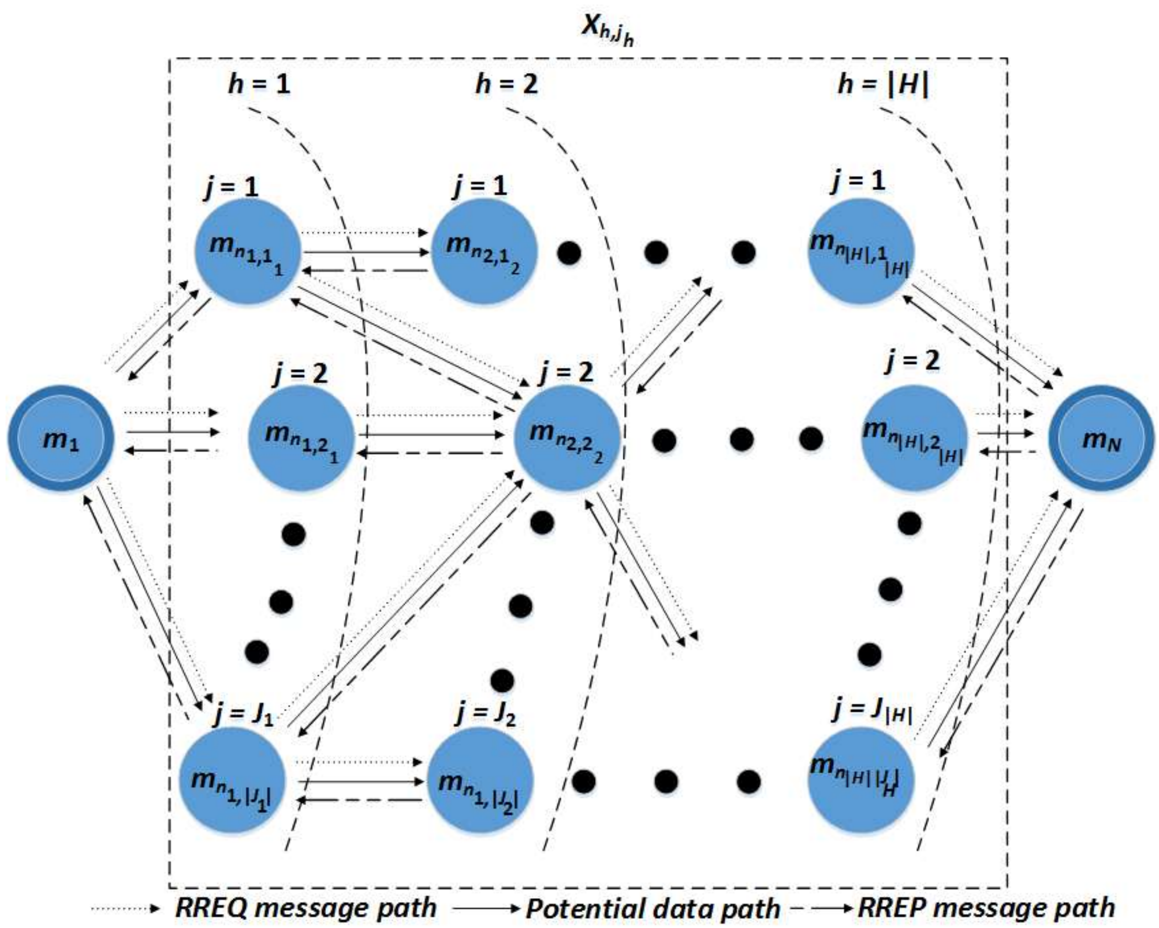


Figure 5-5: A network topology

CR, specifically the dynamicity of the PUs' activities. A cross-layer approach for physical, data link and network layers are left as a future work.

The proposed system model exploit the usage of route request (RREQ) and route reply (RREP) messages, which have been used in traditional routing schemes (e.g., AODV), to broadcast and gather route(s) information. Suppose, a source node m_1 does not have a route information to its destination node m_N in its routing table. It broadcasts route record list (i.e., $\mathbb{R}_{m_1,m_N} = \emptyset$) using a route request RREQ message in the network to discover all possible routes in the network in two main steps. Firstly, the source node m_1 appends its node ID to the route record list $\mathbb{R}_{m_1,m_N} \leftarrow m_1$, which is included in the RREQ message, and broadcasts it to its next-hop neighbor nodes $m_{n_{1,J_1}}$, which are located in the first hop from the source node m_1, using a CCC. Secondly, each neighbor node in the first hop $m_{n_{1,j_1 \in J_1}}$ appends its node ID to the route record list $\mathbb{R}_{m_1,m_N} \leftarrow (m_1) \cup m_{n_{1,j_1}}$, and broadcasts the respective RREQ message to its next-hop neighbor nodes $m_{n_{2,J_2}}$, which are located in the second hop from the source node m_1. The similar RREQ broadcast mechanism is repeated for each next-hop neighbor node in the remaining hops to form a route k_k of $m_1 - m_{n_{1,j_1}} - \cdots - m_N$. Upon receiving a number of RREQ messages from different possible routes K, the destination node m_N generates a route reply RREP message for each route $k_k \in K$ and sends it back towards the source node m_1 via intermediate nodes X_{h,j_h}. The RREP includes a bottleneck link record $\Gamma_{\beta,t}^{k_k}$, which is the average channel available time at the bottleneck link of a route k_k at time t. The SU node i estimates the average channel available time $\varphi_{t,c_c,OFF}^{i,j,k_k}$ of channel $c_c \in C$ on the link of a route k_k connecting itself and its SU neighbor node j, using Equation (5.2) [132]:

$$\varphi_{t,c_c,OFF}^{i,j,k_k} = \frac{\lambda_{c_c,ON}^{p}}{\lambda_{c_c,ON}^{p} + \lambda_{c_c,OFF}^{p}} + \frac{\lambda_{c_c,OFF}^{p}}{\lambda_{c_c,ON}^{p} + \lambda_{c_c,OFF}^{p}} e^{-\left(\lambda_{c_c,ON}^{p} + \lambda_{c_c,OFF}^{p}\right)^{t}} \tag{5.2}$$

Generally speaking, a node i updates the bottleneck link record $\Gamma_{\beta,t}^{k_k}$ if its link to its upstream node j is lower (or $\varphi_{t,c_c,OFF}^{i,j,k_k} < \Gamma_{\beta,t}^{k_k}$). There are two main steps involved in sending the bottleneck link record using the RREP message from a destination node to a source node. Consider a single route $k_k \in K$. Firstly, the destination node m_N initializes the bottleneck link record $\Gamma_{\beta,t}^{k_k}$ with the average channel available time of its link connecting to its upstream neighbor node $m_{n_{|H|,j_{|H|}}} \in X_{h,j_h}$, specifically $\Gamma_{\beta,t}^{k_k} \leftarrow \varphi_{t,c_c,OFF}^{m_N,m_{n_{|H|,j_{|H|}}},k_k}$. The bottleneck link capacity $\Gamma_{\beta,t}^{k_k}$ is included in the RREP message, and it is sent to its upstream neighbor nodes $m_{n_{|H|,j_{|H|}}} \in X_{h,j_h}$ using a CCC. Secondly, the upstream node $m_{n_{|H|,j_{|H|}}}$ updates the bottleneck link record $\Gamma_{\beta,t}^{k_k}$ in the RREP message if the average channel available time of the link connecting to its upstream neighbor node is lower, specifically $\varphi_{t,c_c,OFF}^{m_{n_{|H|,j_{|H|}}},m_{n_{|H|-1,j_{|H|-1}}},k_k} < \Gamma_{\beta,t}^{k_k}$. The remaining nodes in a route k_k follow the same procedure until the RREP message has reached the source node m_1.

5.4.3 Decision making engine for RL-based schemes

This section proposes two RL-based schemes, namely the traditional RL scheme (or TRL henceforth) and a RL scheme with average Q-value (or ARL henceforth), as the decision-making engine for route selection. In general, TRL and ARL share a similar algorithm except the way in which the Q-values are updated: TRL calculates the Q-values using the

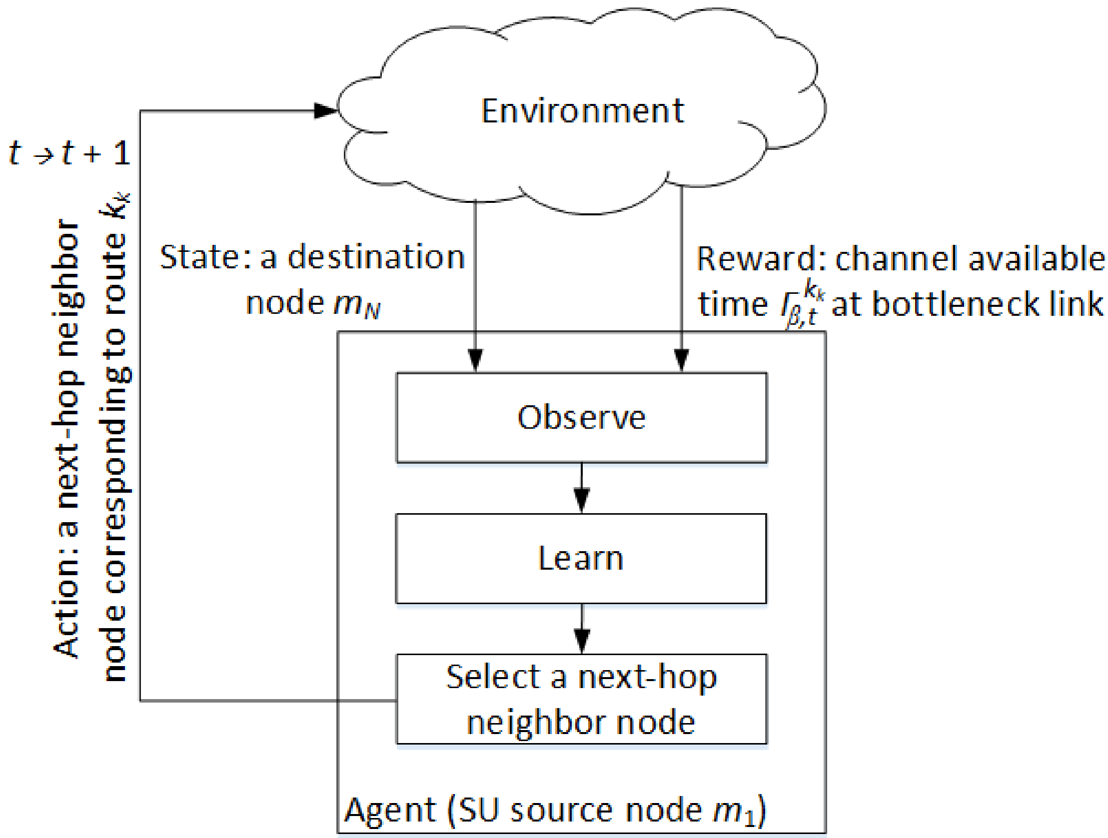


Figure 5-6: Decision making engine for RL-based schemes

traditional approach [133], whereas ARL uses an average Q-value. In general, Q-values constitute knowledge that represents the suitability of an action in a particular state (or operating environment). The decision making engine for the RL-based schemes is shown in Figure 5.6, and it is embedded in a SU source node so that it can select a route in which the average channel available time is the highest possible (or the PUs' activities are minimal) in order to increase throughput and packet delivery ratio, as well as to reduce the number of route breakages. The RL-based scheme uses a distributed model (see Section 2.4), in which PUs do not share spectrum occupancy map with SUs, and so the SUs must sense for available channels and calculate the average channel available time.

The RL agent receives an update of channel state information of the links of each route in the network (i.e., the average channel available time) using RREQ and RREP (see Section 5.4.2). Then, the SU source node m_1 (or the RL agent) selects a route based on Algorithm 1. There are three representations, namely state, action and reward. Generally speaking, the SU source node (agent) selects a neighbor node corresponding to a route k_k (action) to its destination node m_N (state) based on the channel state information (reward). Hence, with m_1 being the source node, the state $s_t^{m_N} \in S$ represents a SU destination node, the action $a_t^{m_{n_{1,j_1}}} \in A = \{m_{n_{1,1_1}}, \ldots, m_{n_{1,|J_1|}}\}$ represents the selection of a neighbor node $m_{n_{1,j_1}}$ of the source node m_1, and the reward R represents the positive or negative consequence of the action taken in the state, which is the highest channel available time $\Gamma_{\beta,t}^{k_k}$ at the bottleneck link of a route k_k connecting a source node and a destination node, and so the reward varies with the dynamicity of the PUs' activities. Table 5-2 shows the RL-based model embedded in the SU source node m_1.

Table 5-2: RL-based model embedded in the SU source node m_1

State	$s_t^{m_N} \in S = \{m_1, m_{n_{1,1_1}}, m_{n_{1,2_1}}, \ldots, m_N\}$, where each state $s_t^{m_N}$ represents a potential SU destination node
Action	$a_t^{m_{n_{1,j_1}}} \in A = \{m_{n_{1,1_1}}, m_{n_{1,2_1}}, \ldots, m_{n_{1,\|J_1\|}} \| m_{n_{1,1_1}} \in k_k\}$, where A is the set of neighboring nodes of source node m_1, and the action $a_t^{m_{n_{1,j_1}}}$ represents the selection of a node $m_{n_{1,j_1}}$, and hence the selection of route $k_k \in K$
Reward	Channel available time $\Gamma_{\beta,t+1}^{k_k}$ at the bottleneck link for a selected route k_k at time $t+1$ between a source node and a destination node

Algorithm 1 shows three steps to select a route in the proposed RL schemes. In Step 1, the RL agent interacts with the operating environment using RREQ and RREP messages to obtain updated information about a set of routes K, including channel state information of the routes in the network (i.e., the average channel available time of each link along a route between a source node m_1 and a destination node m_N). In Step 2, based on the average channel available time, the RL agent computes the Q-value of each route $k_k \in K$ in the network. In Step 3, the RL agent selects a route k_k that offers the highest Q-value. Subsequently, the RL agent uses the selected route k_k for data transmission, and the route consists of multiple links operating on different channels. When the PUs' activities reappear in any of these channels assigned to one of the links along the route k_k, the route is considered broken. This is followed by the transmitting node (or upstream node) of the respective link sending a route breakage message to the source node. The source node selects another route in the next time window $t + 1$.

Algorithm 1: Route selection mechanism with Q-value computation for RL-based schemes at SU node i

/* Step 1: */
/* RREQ messages propagation */
if receive $\mathbb{R}_{m_1,m_N}$ and $i! = m_N$ **then**
  $\mathbb{R}_{m_1,m_N} \leftarrow (\mathbb{R}_{m_1,m_N}) \cup i$
/* RREP message propagation */
else if receive $\mathbb{R}_{m_1,m_N}$ and $i == m_N$ **then**
  Receive RREP for route $k_k \in K$
  for link (i,j) in k_k **do** /* node j is an upstream node of node i */
    Estimate $\varphi_{t,c_c,OFF}^{i,j,k_k}$ using Equation (5.2)
    /* mechanism in a destination node */
    if $i == m_N$
      $\Gamma_{\beta,t}^{k_k} \leftarrow \varphi_{t,c_c,OFF}^{i,j,k_k}$
    /* mechanism in an intermediate nodes and source node */
    else if $(i! = m_1 \mid\mid i! = m_N)$
      if $\varphi_{t,c_c,OFF}^{i,j,k_k} >= \Gamma_{\beta,t}^{k_k}$
        $\Gamma_{\beta,t}^{k_k} \leftarrow \Gamma_{\beta,t-1}^{k_k}$ /* $\Gamma_{\beta,t}^{k_k}$ is not updated */
      else if $\varphi_{t,c_c,OFF}^{i,j,k_k} < \Gamma_{\beta,t}^{k_k}$
        $\Gamma_{\beta,t}^{k_k} \leftarrow \varphi_{t,c_c,OFF}^{i,j,k_k}$ /* $\Gamma_{\beta,t}^{k_k}$ is updated */
      end if
    end if
    if $i! = m_1$
      Send RREP with $\Gamma_{\beta,t}^{k_k}$ to upstream node i
    end if
  end for
end if
/* Step 2: Source node m_1 updates Q-value */
Update Q-value $Q_{t+1}^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$ using $\begin{cases}\text{Equation (5.3) for TRL}\\ \text{Equation (5.5) for ARL}\end{cases}$
/* Step 3: Source node m_1 determines action*/
Determine a_t^* using Equation (5.4)

4.4.3.1 Traditional RL-based scheme

In TRL, the SU source node selects a next-hop neighbor node $a_t^{m_{n_{1,j_1}}}$, which corresponds to a route k_k, leading towards its destination node $s_t^{m_N}$ at time t. It receives its reward in the form of channel available time $\Gamma_{\beta,t+1}^{k_k}$ at the bottleneck link, and updates the corresponding Q-value $Q_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$ for the state-action pair at time $t+1$ as follows:

$$Q_{t+1}^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}) \leftarrow (1-\alpha) \times Q_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}) + \alpha \times \Gamma_{\beta,t+1}^{k_k} \tag{5.3}$$

where $0 \leq \alpha \leq 1$ is the learning rate. When α is higher, the Q-value is more dependent on the current knowledge (or the reward, which is the channel available time $\Gamma_{\beta,t+1}^{k_k}$ at the bottleneck link of route k_k at time $t+1$); and when α is lower, the Q-value is more dependent on the previous knowledge (or the Q-value $Q_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$ at time t). Based on Equation (5.3), the source node m_1 selects the next-hop neighbor node a_t^*, which corresponds to a route k_k, with the highest Q-value as follow:

$$a_t^* = \underset{a \in A}{\operatorname{argmax}}\, Q_t^{m_1}(s_t^{m_N}, a) \tag{5.4}$$

4.4.3.2 RL-based scheme with average Q-value

The average Q-value based route selection scheme (ARL) has been shown to improve stability in simulation setting [134-135]. In this approach, the average Q-value $\bar{Q}_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$ is calculated, and it is used in Q-function $Q_{t+1}^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right)$ to select more stable routes as follows [135]:

$$Q_{t+1}^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right) \leftarrow (1-\alpha) \times Q_t^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right) + \alpha \times \left(\Gamma_{\beta,t+1}^{k_k} + \bar{Q}_t^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right)\right) \tag{5.5}$$

The average Q-value $\bar{Q}_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})$ is a ratio of the sum of all Q-values $Q_{t+1}^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right)$ up to time t to the total number of times if the route k_k is selected for transmission (similarly for the case if the route k_k is not selected), and it is calculated as follows:

$$\bar{Q}_t^{m_1}\left(s_t^{m_N}, a_t^{m_{n_{1,j_1}}}\right) = \begin{cases} \dfrac{\sum_{k_k \epsilon K} P(k^+ = 1|k_k) Q_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})}{\sum_{k_k \epsilon K} P(k^+ = 1|k_k)} & ; k_k \text{ is selected} \\ \dfrac{\sum_{k_k \epsilon K} P(k^- = 0|k_k) Q_t^{m_1}(s_t^{m_N}, a_t^{m_{n_{1,j_1}}})}{\sum_{k_k \epsilon K} P(k^- = 0|k_k)} & ;\ k_k \text{ is not selected} \end{cases} \tag{5.6}$$

where $P(k^+ = 1|k_k)$ is the probability that route k_k is selected for transmission; and similarly, $P(k^- = 0|k_k)$ is the probability that route k_k is not selected for transmission. Also, $\sum_{k_k \epsilon K} P(k^+ = 1|k_k)$ represents the sum of all the probability that route k_k is selected for transmission, which determines the total number of times the route is selected; and similarly, $\sum_{k_k \epsilon K} P(k^- = 0|k_k)$ represents the sum of all the probability that route k_k is not selected for transmission, which determines the total number of times the route is not selected. With the consideration of average Q-value, which is dependent on the probabilities of selecting (or not selecting) a route, the RL agent can select a stable route that has been selected for transmission in the past repeatedly. Based on (5.5), the source node selects the next-hop neighbor node a_t^*, which corresponds to route k_k, with the highest Q-value using Equation (5.4).

4.4.3.3 An illustration of RL-based schemes

Consider an experimental setup to examine the RL-based schemes in a multi-hop CR network as shown in Figure 5.7, which is a simplified topology representation of Figure 5.1. There are six SUs: a source node m_1, intermediate nodes m_2, m_3, m_4, m_5, and a destination node m_6. Initially, the source node m_1 has no route information leading to its destination node m_6 in the route record list (i.e., $\mathbb{R}_{m_1,m_6} = \emptyset$). Then, the source node m_1 appends its own ID to the route record list (i.e., $\mathbb{R}_{m_1,m_6} \leftarrow m_1$) in a newly generated RREQ message, and broadcasts the RREQ message to its neighboring nodes m_2 and m_3 using a CCC to discover routes leading to the destination node m_6. When the neighboring nodes m_2 and m_3 receive separate RREQ messages from the source node m_1, node m_2 appends its ID to the route record list (i.e., $\mathbb{R}_{m_1,m_6} \leftarrow (m_1) \cup m_2$), and node m_3 does the same (i.e., $\mathbb{R}_{m_1,m_6} \leftarrow (m_1) \cup m_3$). Next, nodes m_2 and m_3 forward their respective RREQ messages via CCC to their respective next-hop neighboring nodes m_4 and m_5. The similar procedure is repeated at nodes m_4 and m_5 until the RREQ messages reach destination node m_6. In Figure 5.7, the destination node m_6 receives four possible routes, namely route k_1

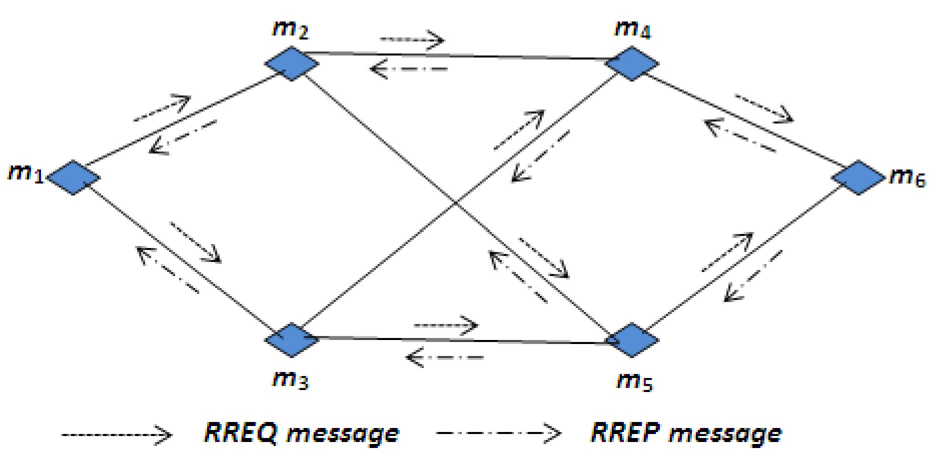


Figure 5-7: A 6-node topology for experimental study with RREQ and RREP message exchanges

$(m_1 - m_2 - m_4 - m_6)$, route k_2 $(m_1 - m_2 - m_5 - m_6)$, route k_3 $(m_1 - m_3 - m_4 - m_6)$ and route k_4 $(m_1 - m_3 - m_5 - m_6)$.

Subsequently, the destination node m_6 generates RREP messages which traverse back towards the source node m_1 using the reversed route given in the RREQ messages. Consider route k_1. The destination node m_6 obtains the average channel available time $\varphi_{t,c_c,OFF}^{m_4,m_6,k_1}$ of the link $m_4 - m_6$ using Equation (5.2) and updates the channel available time of the bottleneck link $\Gamma_{\beta,t}^{k_1}$ in the RREP message. Next, the destination node m_6 sends the RREP message to its upstream node m_4. When node m_4 receives the RREP message, it obtains the average channel available time $\varphi_{t,c_c,OFF}^{m_2,m_4,k_1}$ of the link $m_2 - m_4$. If $\varphi_{t,c_c,OFF}^{m_2,m_4,k_1}$ is smaller than the channel available time of the bottleneck link $\Gamma_{\beta,t}^{k_1}$ in the RREP message (or $\varphi_{t,c_c,OFF}^{m_2,m_4,k_1} < \Gamma_{\beta,t}^{k_1}$), then node m_4 updates the channel available time of the bottleneck link (or $\Gamma_{\beta,t}^{k_1} = \varphi_{t,c_c,OFF}^{m_2,m_{4,k_1}}$) in the RREP message; otherwise the channel available time of the bottleneck link remains the same (or $\Gamma_{\beta,t}^{k_1} = \varphi_{t,c_c,OFF}^{m_4,m_6,k_1}$). The same process is repeated until the RREP message reaches the source node m_1.

Upon receiving RREP messages for the four routes, the source node m_1 computes the Q-value of each route using Equation (5.3) for TRL (or Equation (5.5) for ARL), and selects the route which has the highest Q-value using Equation (5.4). Figure 5.8 shows the trajectory of a selected route on the basis of the highest Q-value at different time instances. For instance, at time instance $t = 0$, route k_3 is selected, as there are no PUs' activities in the respective channels of the links (e.g., channel c_1 is used in link m_1-m_3, channel c_4 in m_3-m_4, and channel c_6 in m_4-m_6) of the route, and so it has the highest Q-value. Whereas,

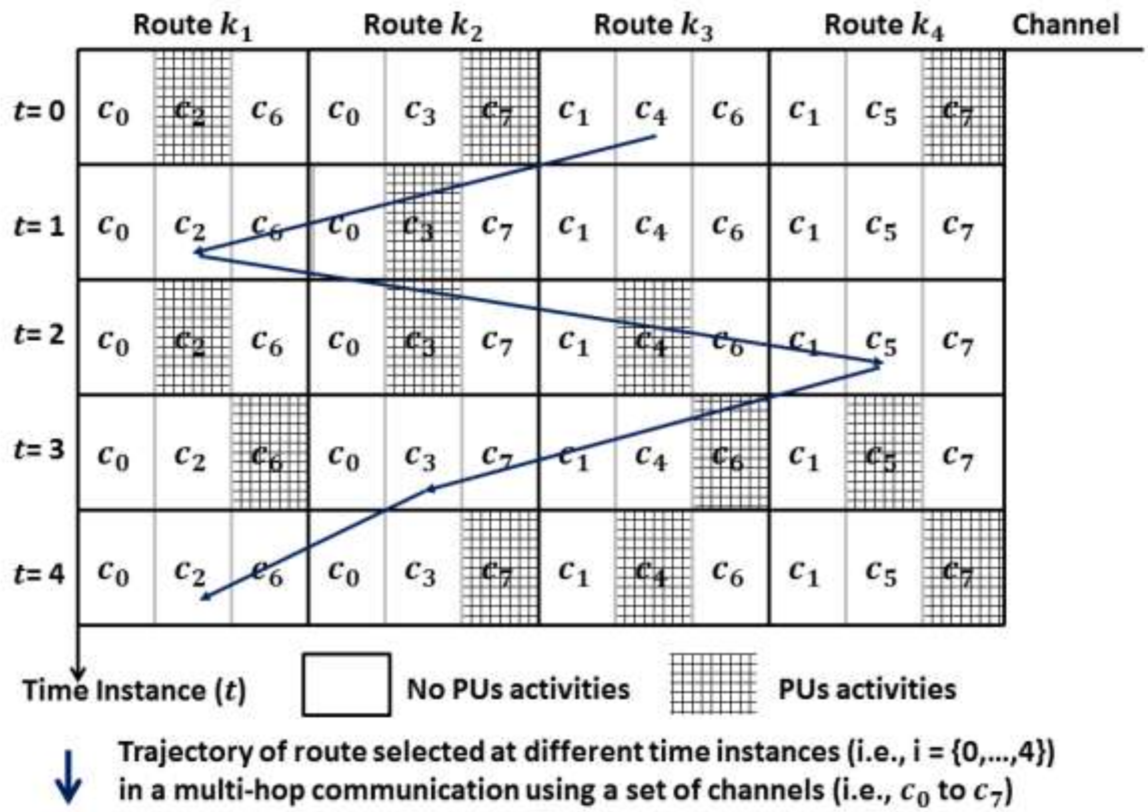


Figure 5-8: Trajectory of route selection at different time instances

route k_1, route k_2, and route k_4 are not chosen due to the presence of PUs' activities in channels c_2 and c_7, which are both chosen by the links $m_2 - m_4$ and $m_5 - m_6$ of those routes.

5.4.4 Decision making engine for SL-based scheme

The decision making engine for SL-based scheme is shown in Figure 5.9, and it is embedded in a SU source node so that it can select the best possible route from a source node to a destination node in order to increase throughput and packet delivery ratio, as well as reduce the number of route breakages. The SL-based scheme uses a centralized model (see section 2.4), in which the PUs share their spectrum occupancy map (i.e., ON duration $\tau_{c_c,ON}^{p}$ and OFF duration $\tau_{c_c,OFF}^{p}$) with SUs located within their transmission range. In

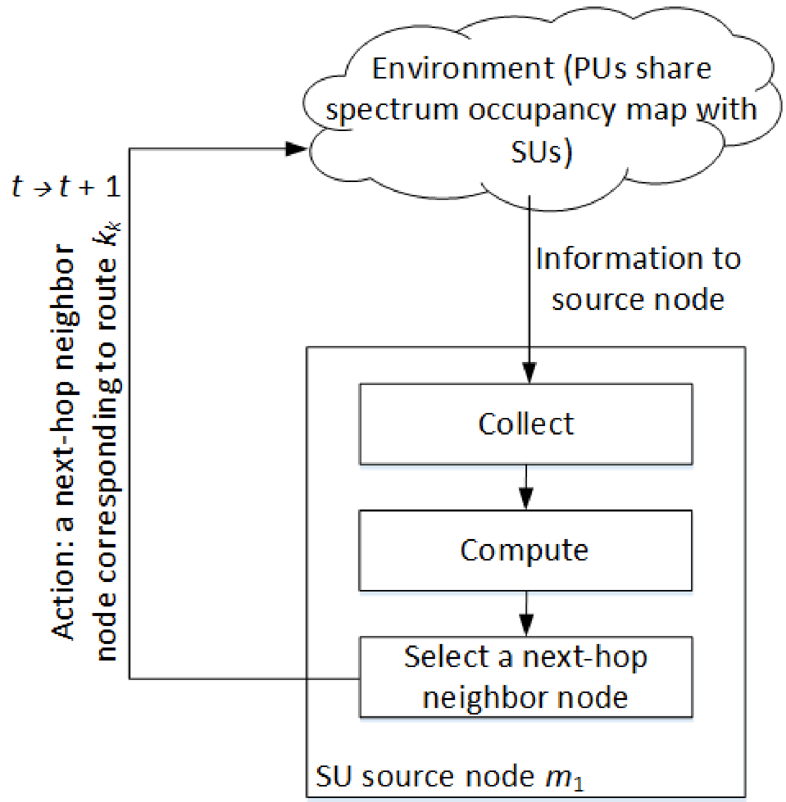


Figure 5-9: Decision making engine for SL-based scheme

practice, the PUs can gain monetary rewards from SUs for sharing the spectrum occupancy map with them. The PUs only allow SUs to use their channels whenever the PUs are in their inactive state (i.e., OFF duration $\tau^{p}_{c_c,OFF}$). It is also beneficial for SUs to be aware of the OFF duration $\tau^{p}_{c_c,OFF}$ time of PUs, which provides them with the exact channel available time (i.e., $e^{i,j,k_k}_{t,c_c,OFF} \leftarrow \tau^{p}_{c_c,OFF}$) for transmission of their packets. For instance, SU node i is in the transmission range of PU p. The PU p shares its $\tau^{p}_{c_c,OFF}$ time, which is the OFF duration of channel c_c, with SU node i. So, the SU node i can use this exact channel available time $e^{i,j,k_k}_{t,c_c,OFF}$ for packet transmission with its neighboring SU node j, which corresponds to route k_k. Algorithm 2 shows three steps to select a route in the proposed SL scheme. In Step 1, every SU node (i.e., $m_1 \in M$ or $X_{h,j_h} \in M$ or $m_N \in M$) receives spectrum occupancy maps (i.e., ON duration $\tau^{p}_{c_c,ON}$ and OFF duration $\tau^{p}_{c_c,OFF}$) from PUs within their respective transmission ranges. In Step 2, the SU source node m_1 uses RREQ and RREP messages to collect the exact channel available time $\mathbb{e}^{i,j,k_k}_{t,c_c,OFF}$ of the links along all possible routes K. Based on $\mathbb{e}^{i,j,k_k}_{t,c_c,OFF}$, the SU source node computes the bottleneck link $\Gamma^{k_k}_{\beta,t}$ of each route $k_k \in K$ in the network. In Step 3, the SU source node selects a route k_k that offers the highest channel available time at its bottleneck link. Subsequently, the source node uses the selected route k_k for data transmission, and the route consists of multiple links operating on different channels. When the PUs' activities reappear in any of these channels assigned to one of the links along the route k_k, the route is considered broken. This is followed by the transmitting node (or upstream node) of the respective link sending a route breakage message to the source node. The source node selects another route in the next time window $t + 1$.

In contrast to the RL-based decision making engine, which is embedded in the SU source node only, the SL-based decision making engine is embedded in each SU node. The PUs provide their spectrum occupancy map (i.e., ON duration $\tau^{p}_{c_c,ON}$ and OFF duration $\tau^{p}_{c_c,OFF}$ for their respective channel c_c) to SU nodes that are within their transmission range. The SU source node m_1 receives these updates using RREQ and RREP. Upon receiving the exact channel available time $\mathbb{e}^{i,j,k_k}_{t,c_c,OFF}$ of every link (e.g., a link between SU node i and SU node j is m_i-m_j), the SU source node m_1 obtains the minimum channel available time $\Gamma^{k_k}_{\beta,t}$ at the bottleneck link along the route k_k from source node m_1 to destination node m_N, which can be computed as follows:

$$\Gamma_{\beta,t}^{k_k} = \underset{(i,j)\in k}{argmin} ⅇ_{t,c_c,OFF}^{i,j,k_k} \qquad \forall\, k_k \in K \tag{5.7}$$

Based on Equation (5.7), the source node selects the next-hop neighbor node $a \in m_{n_{1,j_1}}$, which corresponds to route k_k, that offers the highest minimum channel available time $\Gamma_{\beta,t}^{k_k}$ at its bottleneck link as follows:

$$a_t^* = \underset{k_k\in K}{argmax}\, \Gamma_{\beta,t}^{k_k} \tag{5.8}$$

4.4.4.1 *An illustration of SL-based scheme*

Consider an experimental setup to examine the SL-based scheme in a multi-hop CR network as shown in Figure 5.7. The SL-based scheme uses a centralized model, in which the PUs share their respective spectrum occupancy map (i.e., ON duration $\tau_{c_c,ON}^{p}$ and OFF duration $\tau_{c_c,OFF}^{p}$) with SUs located within their respective transmission range. So, the source node m_1 needs to collect information about the routes and the exact channel access time $ⅇ_{t,c_c,OFF}^{i,j,k_k}$ of the links in the network. The SL-based scheme shares similar mechanism with the RL-based scheme. The only exception is that, in the SL-based scheme, the SUs receive the exact channel access time rather than the estimated average channel available time $\varphi_{t,c_c,OFF}^{i,j,k_k}$ (see Equation (5.2)). Next, upon receiving RREP messages for the four routes, the source node m_1 obtains the minimum channel available time of the bottleneck link of each route, and selects the route with the highest minimum channel available time using Equation (5.8).

Algorithm 2: Route selection mechanism for the SL-based scheme at node i

/* Step 1 */
Receive ($\mathbb{e}_{t,c_c,OFF}^{i,j,k_k} \leftarrow \tau_{c_c,OFF}^{p}$) from PUs
/* Step 2 */
/* Source node m_1 initiates the RREQ messages propagation */
if receive $\mathbb{R}_{m_1,m_N}$ and $i! = m_N$ **then**
    $\mathbb{R}_{m_1,m_N} \leftarrow (\mathbb{R}_{m_1,m_N}) \cup i$
/* RREP message propagation */
else if receive $\mathbb{R}_{m_1,m_N}$ and $i == m_N$ **then**
    Receive RREP for route $k_k \in K$
    for link (i, j) in k_k do /* node j is an upstream node of node i */
        /* mechanism in a destination node */
        if $i == m_N$
            $\Gamma_{\beta,t}^{k_k} \leftarrow \mathbb{e}_{t,c_c,OFF}^{i,j,k_k}$
        /* mechanism in an intermediate node and source node */
        else if ($i! = m_1 \parallel i! = m_N$))
            if $\mathbb{e}_{t,c_c,OFF}^{i,j,k_k} >= \Gamma_{\beta,t}^{k_k}$
                $\Gamma_{\beta,t}^{k_k} \leftarrow \Gamma_{\beta,t-1}^{k_k}$
            else if $\mathbb{e}_{t,c_c,OFF}^{i,j,k_k} < \Gamma_{\beta,t}^{k_k}$
                $\Gamma_{\beta,t}^{k_k} \leftarrow \mathbb{e}_{t,c_c,OFF}^{i,j,k_k}$
            end if
        end if
        if $i! = m_1$
            Send RREP with $\Gamma_{\beta,t}^{k_k}$ to upstream node i
        end if
    end for
/* Step 3 */
Determine a_t^* using Equation (5.8)

5.5 Experiment and evaluation

This section presents experimental setup and performance evaluation. The experimental parameters for both topologies are shown in Table 5-3.

Table 5-3: Experimental parameters

Category	Adjustable Parameter	Value	
		6-node Topology	10-node Topology
PU	Number of PUs	4	6
	PUs' activities model	Exponential ON-OFF model	
	$\lambda^{p}_{c_c,ON}$	15 seconds	
	$\lambda^{p}_{c_c,OFF}$	{10,20, 30, 40, 50, 60, 70, 80} seconds	
SU	Number of USRP SU nodes	6	10
	Channel sensing time window t_s	≪1s	
	Data transmission time window t_d	3s	
Antenna	Carrier frequency range	824 MHz – 960 MHz	
Network (USRP)	Number of channels	8	13
	Modulation type	GMSK	
	Supported bandwidth	~40 MHz	
	Throughput	8 Mbps	
	Transport layer	UDP	
	Experiment duration	300 seconds	
RL	Learning rate α	{0.1, 0.3, 0.5, 0.7, 0.9}	

5.5.1 Experimental setup

Two experimental scenarios, namely a 6-node topology (see Figure 5.1) and a 10-node topology (see Figure 5.11), for multi-hop CR networks are considered. The 6-node topology and 10-node topology have six and ten USRP SU nodes, respectively. This chapter deploys topologies of up to 10 USRP SU nodes, which extend existing implementations [26-27] with more nodes. Figure 5.10 shows the physical deployment of a 10-node topology in which the USRP SU nodes are connected via a gigabit Ethernet switch to a computer, which runs the GNU radio software that loads the Python program into the USRP SU nodes (see Section 5.3.1). The USRP/ GNU Radio testbed is setup in an indoor environment (i.e., a hall with concrete walls) where the USRP SU nodes are placed on a 2 feet × 3.8 feet table. In Figure 5.10, based on the assumptions (see Section 5.4.2), the USRP SU nodes are placed close to each other with a maximum distance of 4.5 inch between a pair of USRP SU nodes while emulating the main characteristic (i.e., dynamicity of PUs' activities) of a CR environment. In addition, the neighboring links of SUs use

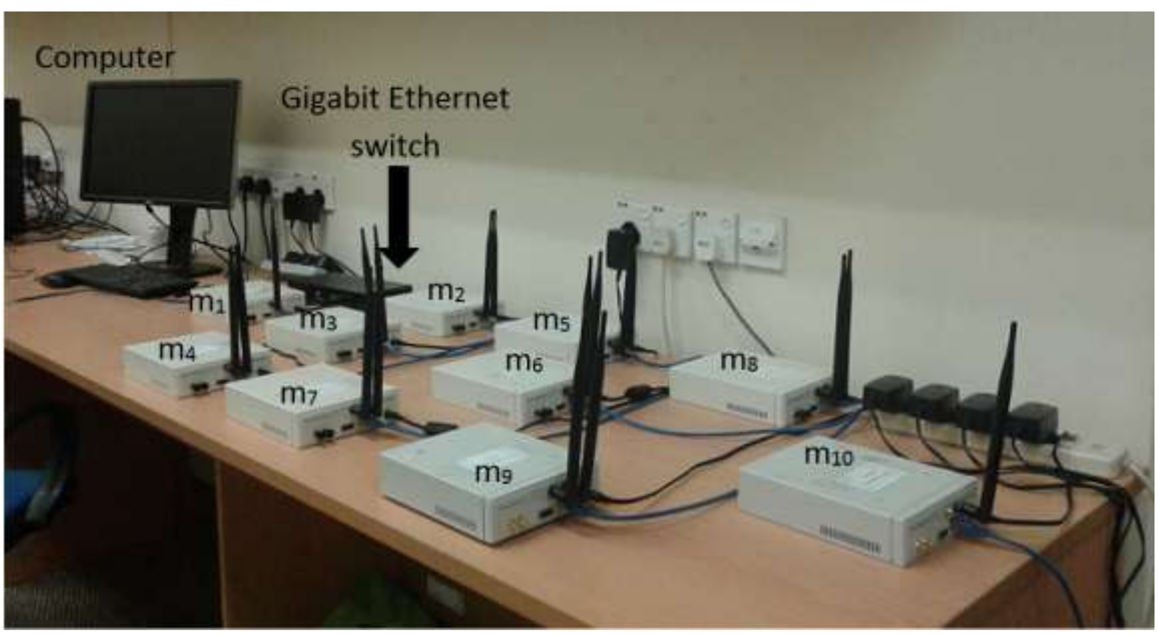


Figure 5-10: Physical deployment of a 10-node topology

distinct channels in order to avoid data link-layer interference among the respective SUs. In this regard, each transceiver uses two distinct frequencies for transmission and reception, and a guard-band of 8 MHz is used in between the two frequencies in order to avoid interference. The computer runs a media player application (i.e., VLC) in a server and client mode and feeds the video into a USRP SU source node. User Datagram Protocol (UDP) is used so that the effects of congestion window in Transmission Control Protocol (TCP) are not considered. The PUs' activities are emulated using an exponential ON-OFF model (see Section 5.4.2) within a Python code. Throughout the experiment, the rate of the ON time duration of PU p in each of its channel $c_c \in C$ is a constant $\lambda^{p}_{c_c,ON} = 15$ s, and the rate of the OFF time duration of PU p in each of its channel $c_c \in C$ is a variable $\lambda^{p}_{c_c,OFF}$ ranging from 10 s to 80 s [26-27]. This means that the channel utilization of PU ranges from 16% (or 15/95) to 60% (or 15/25). The channels with a channel utilization of PU of more than 60% are not considered in this experiment as the SUs can highly interfere with the PUs. Each USRP SU node is equipped with VERT900 antennas [136], and they can operate in frequency ranging from 824 MHz to 960 MHz. As temporal variability occurs in the real-world wireless environment, each experiment is repeated 15 times, and each experiment runs for a duration of 300 s.

The two topologies are selected in order to analyze the QoS performance of the proposed schemes, and to investigate the scalability of the network with 6 and 10 USRP SU nodes. In the 6-node topology as shown in Figure 5.1, there are 6 USRP SU nodes (i.e., a source node m_1, intermediate nodes m_2, m_3, m_4, m_5, and a destination node m_6). There are 4 PUs (see section 5.4.2 for the model of PUs' activities), and each of them

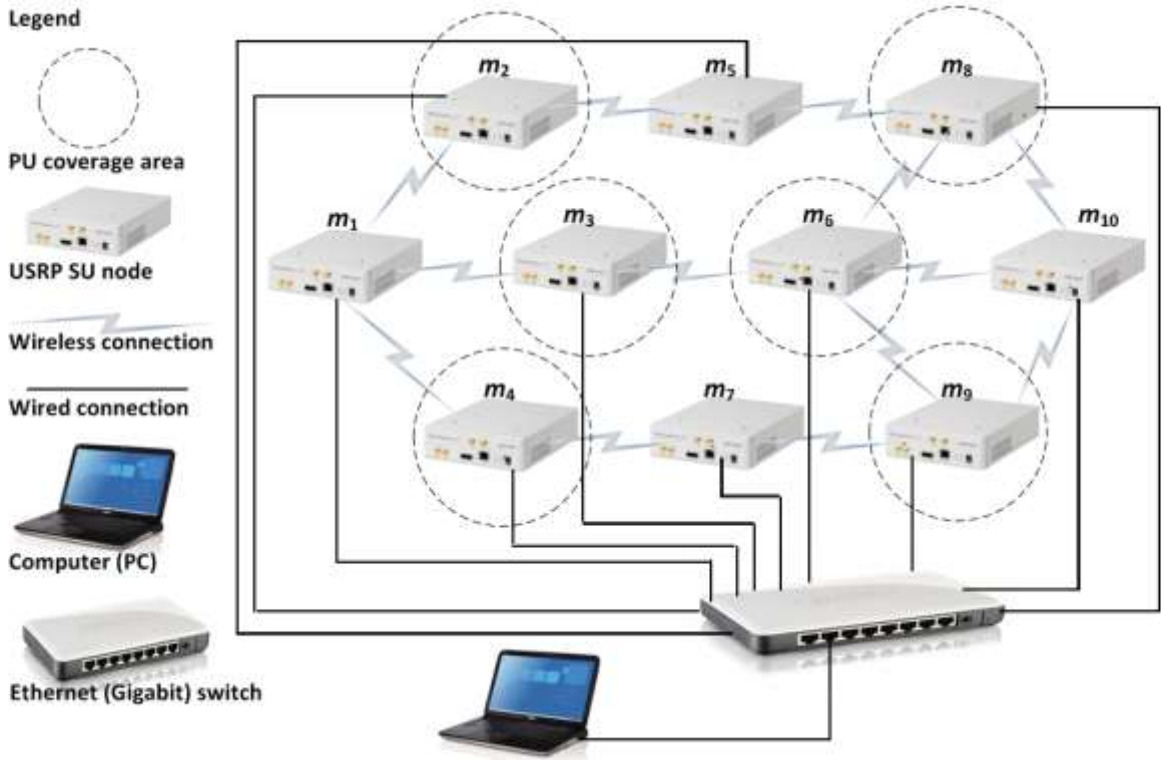


Figure 5-11. A 10-node topology consists of ten USRP SU nodes

interferes with the intermediate nodes (i.e., m_2, m_3, m_4, m_5). Hence, each PU can interfere with 3 SU links from each of the SUs within the PU's coverage area. There are four possible routes from the source node m_1 to the destination node m_6, namely route k_1 $(m_1 - m_2 - m_4 - m_6)$, route k_2 $(m_1 - m_2 - m_5 - m_6)$, route k_3 $(m_1 - m_3 - m_4 - m_6)$ and route k_4 $(m_1 - m_3 - m_5 - m_6)$. In the 10-node topology as shown in Figure 5.11, there are 10 USRP SU nodes: a source node m_1, intermediate nodes $m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9$, and a destination node m_{10}. There are 6 PUs that interfere with the intermediate nodes (i.e., $m_2, m_3, m_4, m_6, m_8, m_9$). At nodes m_2, m_3 and m_4 a PU can interfere with 2 SU links from each of the SUs that lie within the PU's coverage area; while at m_8 and m_9, a PU can interfere with 3 SU links from each of the SUs; and at m_6, a PU can interfere with 4 SU links from each of the SUs. In both topologies, the number of PUs constitutes approximately 60% of the total number of SU nodes. There are five possible routes from the source node m_1 to the destination node m_{10}, namely route k_1 ($m_1 - m_2 - m_5 - m_8 - m_{10}$), route k_2 ($m_1 - m_3 - m_6 - m_8 - m_{10}$), route k_3 ($m_1 - m_3 - m_6 - m_{10}$), route k_4 ($m_1 - m_3 - m_6 - m_9 - m_{10}$) and route k_5 ($m_1 - m_4 - m_7 - m_9 - m_{10}$). Since SU-SU interference is not considered in this investigation based on the assumption (see Section 5.4.2 for explanation), neighboring links between a pair of SUs use distinctive channels in order to avoid interference among the respective SUs. This experiment uses 8 and 13 channels in the 6-node topology and 10-node topology, respectively (see Section 5.4.2). The guard-band between two channels is set to 8 MHz for smooth video transmission and to avoid inter-channel interference.

5.5.2 Performance evaluation

This subsection presents performance evaluation including experiment ordinates, performance metrics, complexity analysis and results.

5.5.2.1 Experiment ordinates and performance metrics

The experiment ordinate is the PUs' OFF time. The PUs' OFF time $\lambda^{p}_{c_c,OFF}$ is the time duration in a time window during which the PUs are inactive. The performance metrics are packet delivery ratio, the number of route breakages, and throughput. The packet delivery ratio is the total number of packets received by the destination node to the total number of packets sent by the source node. A route breakage happens whenever a PU reappears in the channel of any of the links in the route for packet transmission from a source node to a destination node. Lastly, the throughput represents the effectiveness of the network in delivering data packets from a source node to a destination node, and it is measured in bits per second (bps).

5.5.2.2 Complexity analysis

This section presents the complexity analysis of proposed RL- and SL-based schemes in terms of message and time complexities. The message complexity $\mathcal{M}$ is defined as the number of messages exchanged in the network in order to obtain updated information (i.e., channel state information of the routes) about a set of routes K. The channel state information consists of the average channel available time of each link along a route from a source node m_1 to a destination node m_N. When a SU sends a message to each of its neighboring SUs, a single message is incurred, and so the message complexity $\mathcal{M}$ is increased by one. The time complexity T is defined as the number of time steps incurred to perform route selection, which covers finding the number of available routes in the network, selecting a route and switching from a broken route, which may occur due to the re-appearance of the PUs' activities at the bottleneck link of the route to another one. To calculate the time complexity discrete time steps are considered. One time step is the time incurred between the transmission of a message from a SU sender node and the reception of the message at its SU receiver node. Suppose, a SU source node generates a RREQ message and broadcast it towards its neighboring SU nodes. This process continues until the message reaches its SU destination node. Denote the average number of neighbor nodes for each node by η_i, and the intermediate nodes are up to h hops away from the SU source node. So, the whole process of RREQ message propagation takes $\mathcal{M}_{RREQ} = \eta_i \times (h+1)$ messages and $T_{RREQ} = h + 1$ time steps. Upon receiving RREQ message, the SU destination node generates RREP message and sends it back on the reverse route $k_k \epsilon K$ that the RREQ message has traversed. The number of intermediate nodes can be denoted from a SU source node to a SU destination node along a route k_k involved in RREQ is equal to the number of hops h. So, the whole process of RREP message propagation takes $\mathcal{M}_{RREP} = \sum_{k_k \in K} (h+1)^{k_k}$ messages and $T_{RREP} = h+1$ time steps. So, the total of message complexity in the proposed schemes is $\mathcal{M} = \mathcal{M}_{RREQ} + \mathcal{M}_{RREP} = \eta_i(h+1) + \sum_{k_k \in K}(h+1)^{k_k})$ and the time complexity is $T = T_{RREQ} + T_{RREP} = 2(h+1)$ time steps.

5.5.2.3 Experimental results

This section presents experimental results. Section 5.5.2.3.1 presents the results of the effects of learning rate α on the RL scheme. Section 5.5.2.3.2 presents the results of the comparison of different route selection schemes, namely Highest-Channel (HC), RL-based, as well as SL-based schemes. Whereas, section 5.5.2.3.3 presents the comparison of performance between 6-node and 10-node topologies.

5.5.2.3.1 Effects of learning rate α on RL-based schemes

This section presents the effects of learning rate α on the QoS parameters (i.e. throughput and packet delivery ratio, as well as routing stability) of a RL-based scheme (i.e. TRL approach) in the 6-node topology and 10-node topology, respectively. As TRL and ARL schemes produce approximately similar results, only results of the TRL scheme are presented (see section 5.5.2.3.2). The learning rate α is an important parameter that affects the learning speed. In this work, the learning rate α is dependent and adjusted according to the level of dynamicity of the operating environment. Specifically, higher (or lower) α value is needed for operating environment with higher (or lower) dynamicity. As shown in

Figure 5.12 in which higher α value shows greater performance enhancement providing higher throughput compared to lower α value as the PUs' OFF time increases beyond 30 s, with the optimal throughput being achieved with $\alpha = 0.9$.

Figure 5.12 and Figure 5.13 show that the throughput and packet delivery ratio slightly increase (note that the y-axis starts at 1.7 Mbps and 0.88 in the figures, respectively) with increasing learning rate (i.e. $\alpha \geq 0.5$). Similarly, Figure 5.14 shows that the number of route breakages is lesser when the learning rate is higher (i.e. $\alpha \geq 0.5$). This is because, as α increases, the RL-based scheme is more dependent on the current knowledge (i.e., $\Gamma_{\beta,t+1}^{k_k}$) due to the high temporal variability of the wireless channels (i.e., channel available time), rather than the previous knowledge (i.e., $Q_t^{m_{1,k_k}}(s_t, a_t^{m_{n_{1,j_1}}})$). With learning rate α=0.9, the RL approach provides the best possible network performance, and so this value is chosen for comparison with the other approaches in Section 5.5.2.3.2, as well as comparison in the performance achieved by both 6-node and 10-node topologies in Section 5.5.2.3.3.

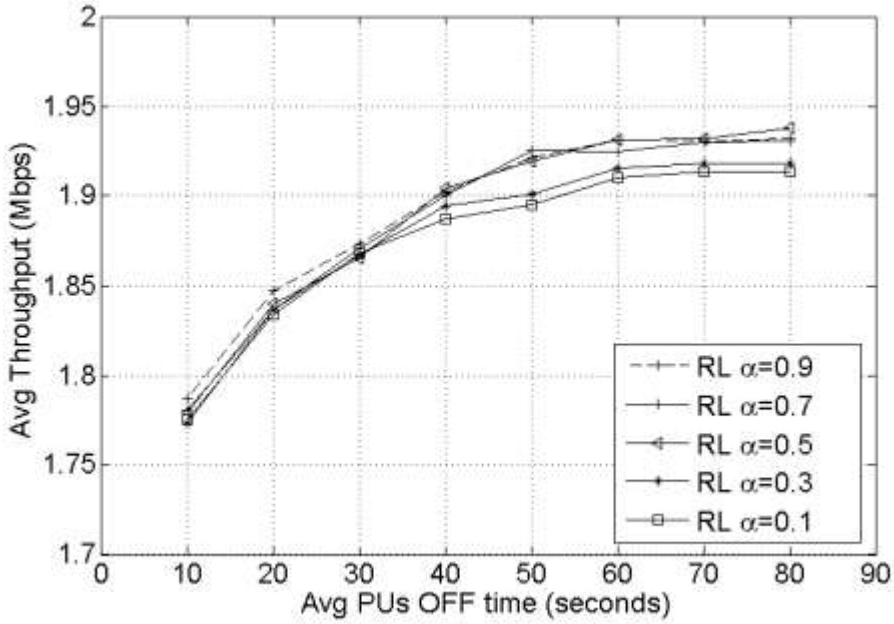


Figure 5-12: Average throughput versus PU-OFF time $\lambda_{c,OFF}^{p}$ at different values of α for TRL scheme using 6-node topology.

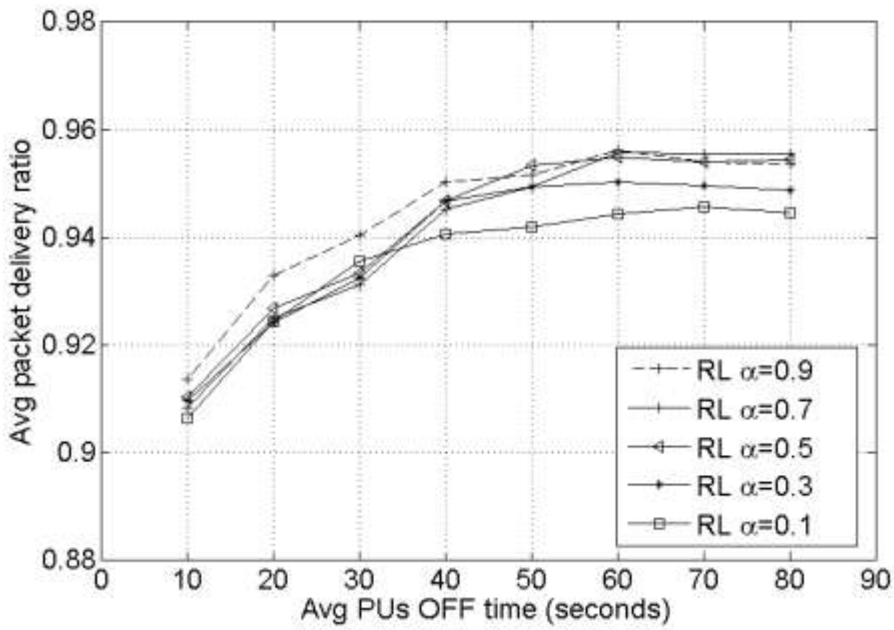


Figure 5-13: Average packet delivery ratio versus PU-OFF time $\lambda_{c,OFF}^{p}$ at different values of α for TRL scheme using 6-node topology.

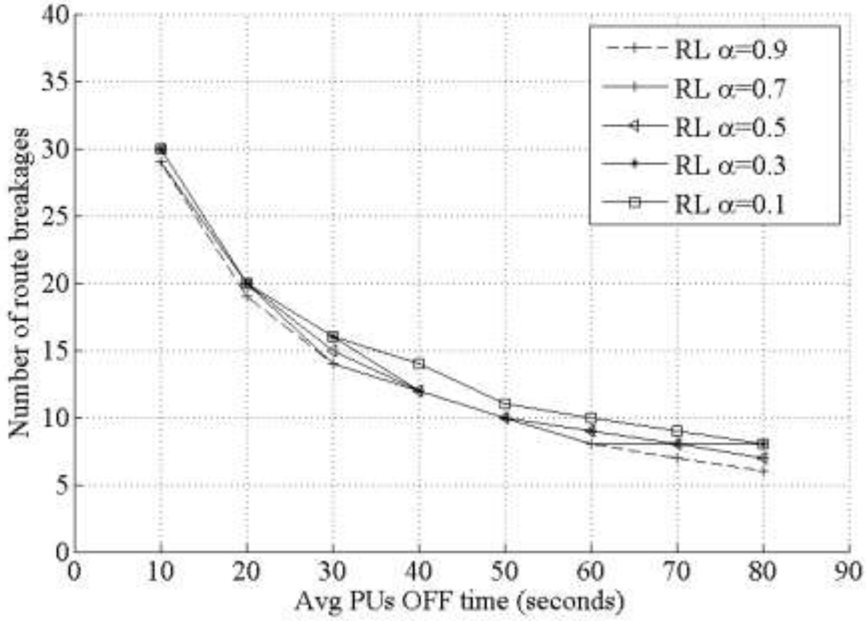


Figure 5-14: Average number of route breakages versus PU-OFF time $\lambda^{p}_{c,OFF}$ at different values of α for TRL scheme using 6-node topology.

While considering the 10-node topology, the effects of learning rate α on RL-based schemes produce similar results, and the graphs of throughput, packet delivery ratio and number of route breakages are shown in Figure 5.15, Figure 5.16 and Figure 5.17, respectively.

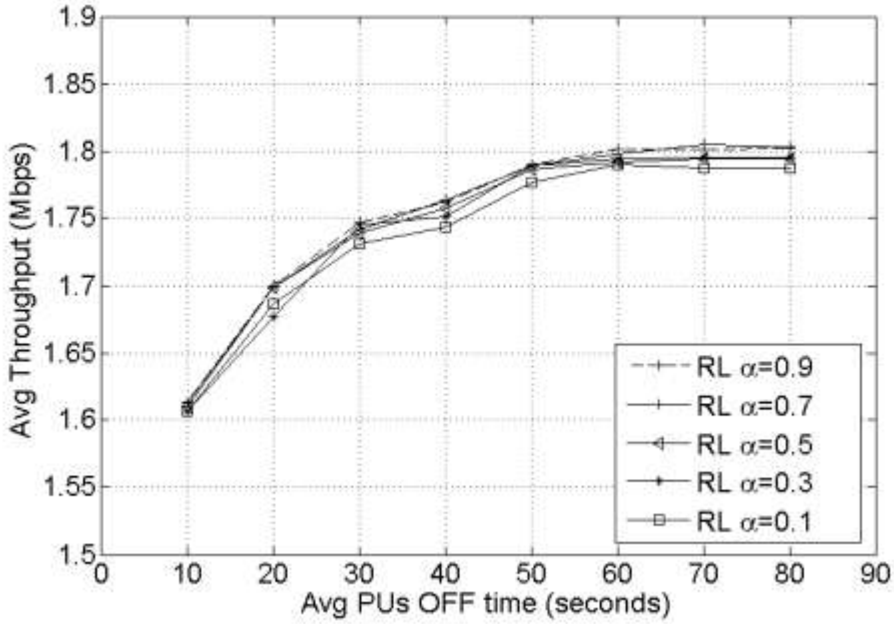


Figure 5-15: Average throughput versus PU-OFF time $\lambda^{p}_{c,OFF}$ at different values of α for TRL scheme using 10-node topology.

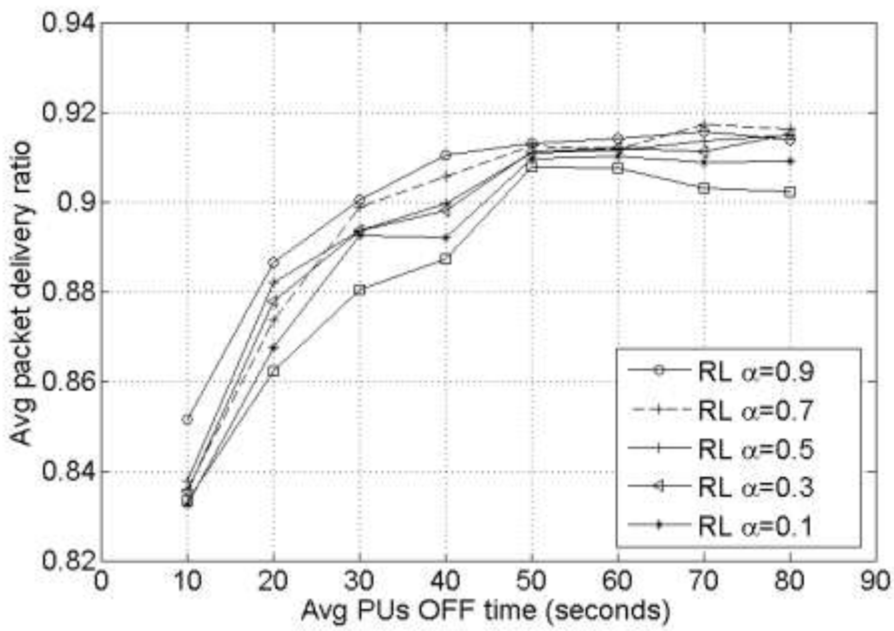


Figure 5-16: Average packet delivery ratio versus PU-OFF time $\lambda^{p}_{c,OFF}$ at different values of α for TRL scheme using 10-node topology.

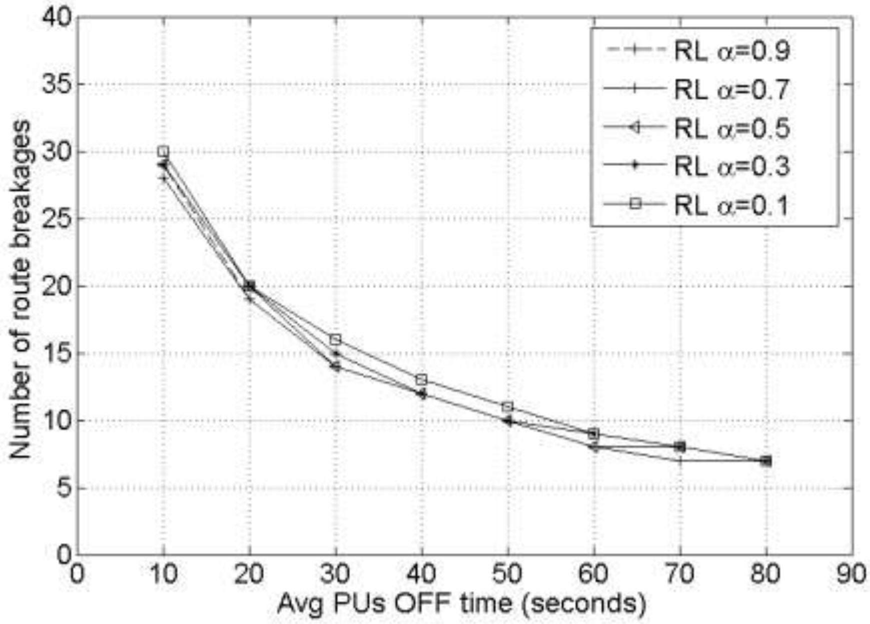


Figure 5-17: Average number of route breakages versus PU-OFF time $\lambda^{p}_{c,OFF}$ at different values of α for TRL scheme using 10-node topology.

5.5.2.3.2 *Comparison of route selection schemes*

This section presents the experimental results of the RL-based schemes (see Section 5.4.3) and SL-based scheme (see Section 5.4.4) in the 6-node and 10-node topologies. However, it is improbable to compare the existing schemes with the proposed RL- and SL- based schemes. This is due to the fact that the system architectures in the existing schemes use additional computational resources to implement the network layer implementation, while this proposed work has only a single computer to implement the network layer implementation (see Section 5.3). The results of the proposed schemes are compared with Highest-Channel (HC), which is a RL-based scheme that selects the route in a multi-hop CR networks with the highest number of available channels [137], instead of highest channel available time. The main objective of the proposed schemes is to select the best possible route from a source node to a destination node in a multi-hop CR networks in order to improve QoS parameters, particularly throughput and packet delivery ratio, as well as the number of route breakages which affects the routing stability. Figure 5.18 and Figure 5.19 show that throughput and packet delivery ratio performance increase with the average PUs' OFF time from 10 s to 50 s in the 6-node topology and stabilize when the average PUs' OFF time reaches approximately 50s. Figure 5.20 shows that the number of route breakages reduces with increasing average PUs' OFF time and stabilizes when the PUs' OFF time reaches 60s. Similarly, Figure 5.21 and Figure 5.22 show that throughput and packet delivery ratio performance increase with the average PUs' OFF time from 10 s to 50 s in the 10-node topology and stabilizes when the average PUs' OFF time reaches approximately 50 s. Figure 5.23 shows that the number of route breakages reduces with increasing average PUs' OFF time and stabilizes when the PUs' OFF time reaches 60 s.

Overall, the SL-based scheme achieves higher throughput and packet delivery ratio, as well as lower number of route breakages, in comparison with the RL-based and HC schemes. This is because the SL-based scheme, SUs are aware of the exact channel available time (or PUs' activities) as PUs share their respective spectrum occupancy map (i.e., $\lambda^{p}_{c_c,OFF}$) with the SUs located within their respective transmission range. Although the RL-based and HC schemes receive ideal sensing outcomes, their performance degrades in comparison to the SL approach due to the channel sensing delay. The number of route breakages of the SL-based and RL-based schemes are lower than that of the HC scheme. In all cases, the HC scheme shows the least performance in comparison with SL-based and RL-based schemes as it selects a route with the highest number of available channels, which may have low channel available time at its bottleneck link. In addition, the ARL scheme shows a very minor improvement in comparison with the TRL scheme for both 6-

node and 10-node topologies. The RL schemes are primarily dependent on the estimated channel available time in order to compute the Q-value. However, the key difference between the TRL scheme and the ARL scheme is that the ARL scheme further consider the average Q-value for the computation of Q-value (see Section 5.4.3); which determines the stability in terms of the route(s) selection in the past.

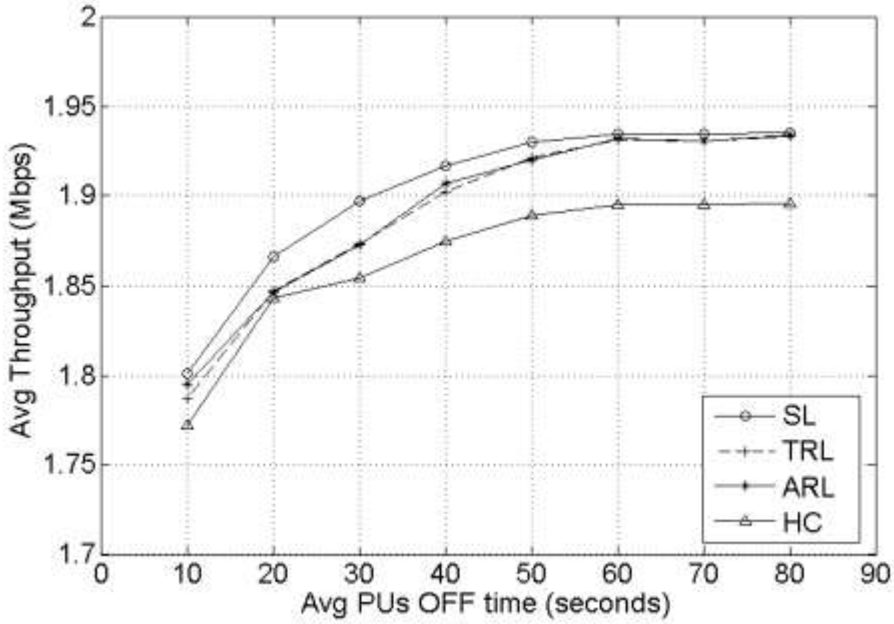


Figure 5-18: Average throughput versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 6-node topology

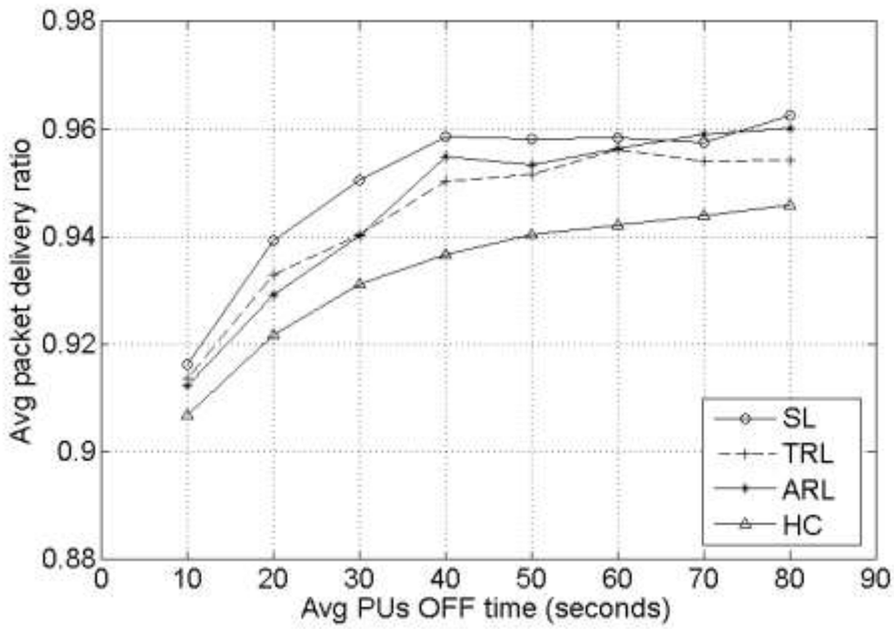


Figure 5-19: Average packet delivery ratio versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 6-node topology

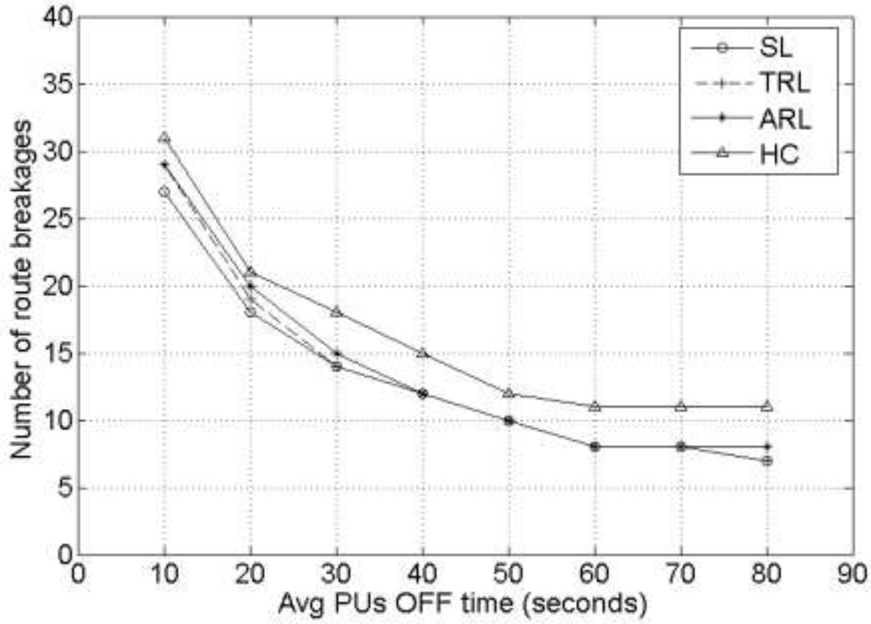


Figure 5-20: Average number of route breakages versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 6-node topology

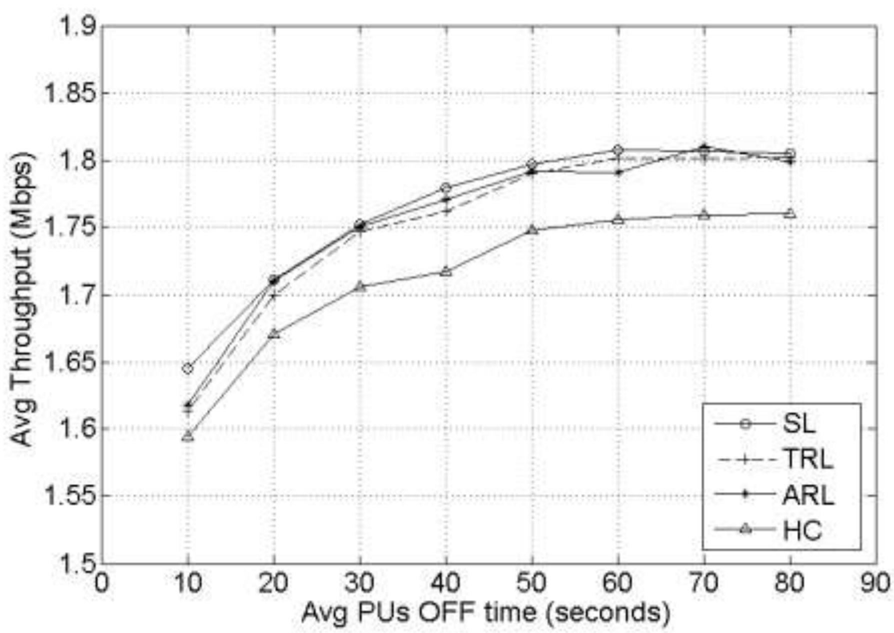


Figure 5-21: Average throughput versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 10-node topology

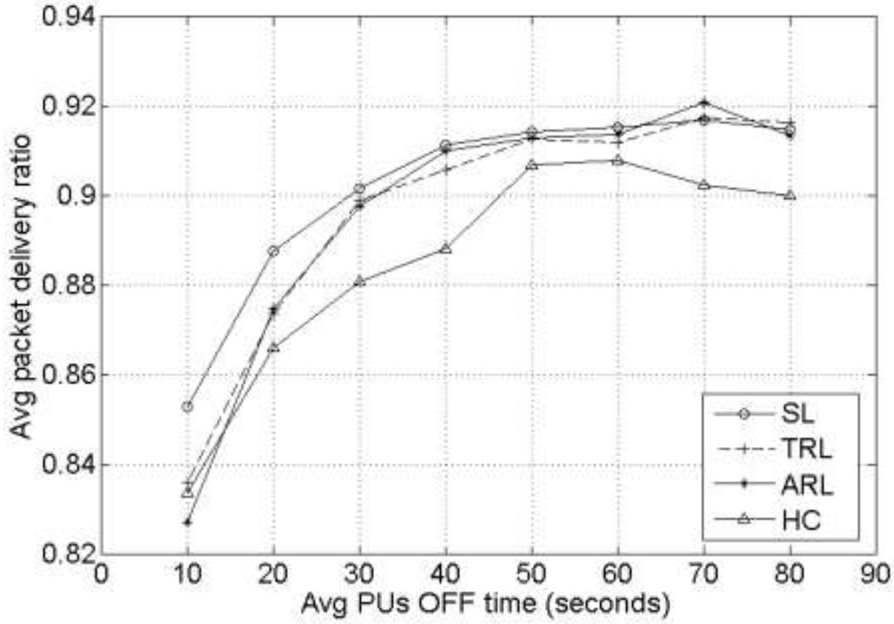


Figure 5-22: Average packet delivery ratio versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 10-node topology.

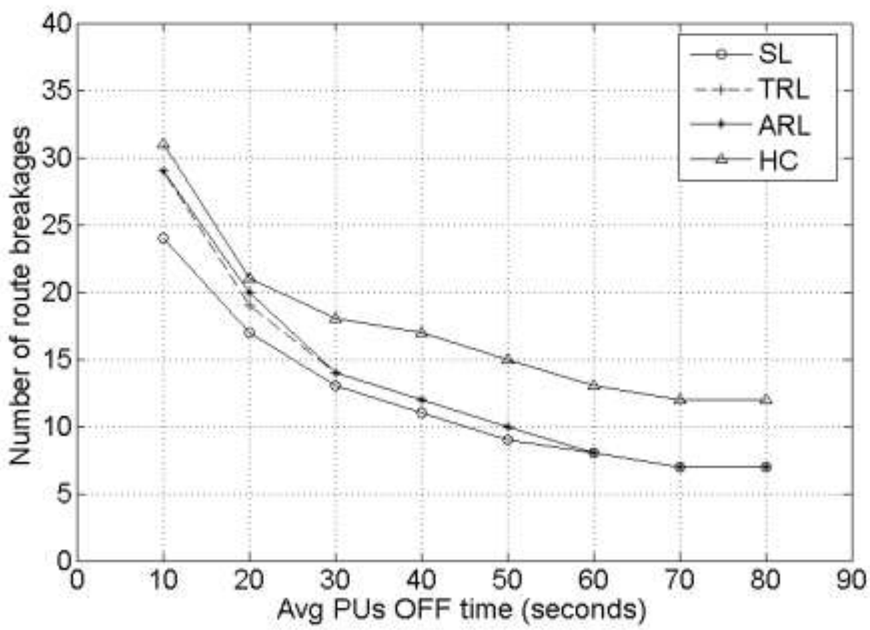


Figure 5-23: Average number of route breakages versus PU-OFF time $\lambda_{c,OFF}^{p}$ for a 10-node topology.

5.5.2.3.3 Comparison of performance between 6-node and 10-node topologies

This section compares the QoS performance (i.e., throughput and packet delivery ratio, as well as routing stability) achieved by TRL in 6-node and 10-node topologies. In general, the 6-node topology provides better network performance compared to the 10-node topology. The maximum throughput achieved by the 6-node and 10-node topologies are 1.93 Mbps and 1.8 Mbps respectively, as shown in Figure 5.24. The maximum packet delivery ratio achieved by the 6-node and 10-node topologies are 95.8% and 91.9% respectively, as shown in Figure 5.25. The performance deteriorates in the 10-node topology due to the fact that it has higher number of hops in a route resulting in higher packet loss as compared to 6-node topology. The number of route breakages are approximately similar in both topologies because of the same PUs activity level (i.e., $\lambda^{p}_{c_c,OFF}$ = {10, 20, 30, 40, 50, 60, 70, 80} s and $\lambda^{p}_{c_c,OFF}$ = 15 s) as shown in Figure 5.26.

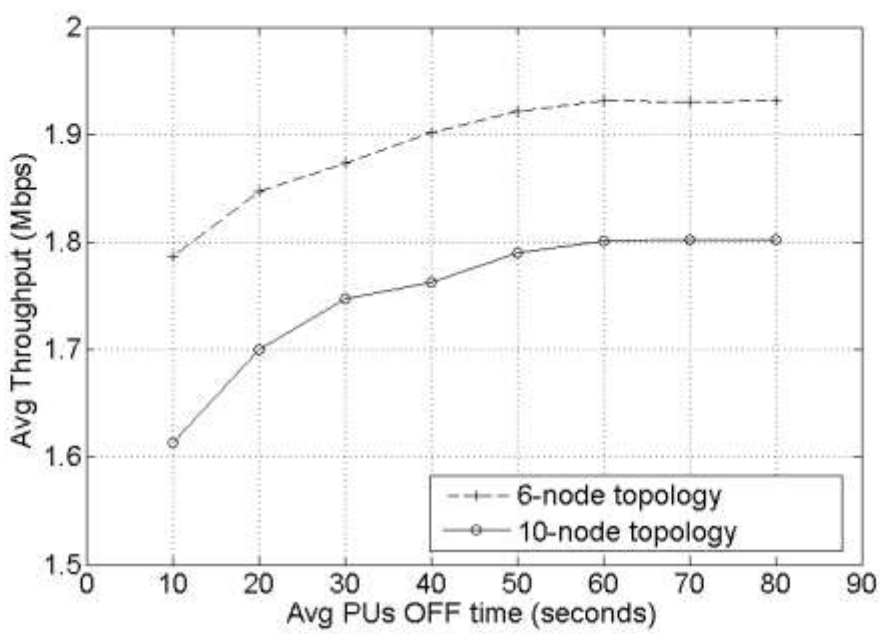


Figure 5-24: Average throughput versus PU-OFF time $\lambda^{p}_{c,OFF}$ at α = 0.9 for TRL scheme in 6-node and 10-node topologies.

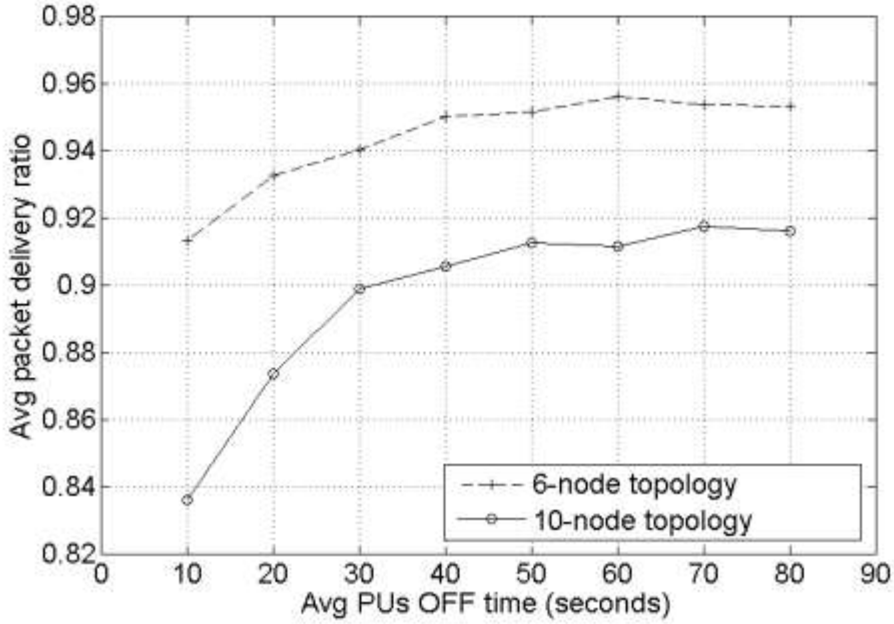


Figure 5-25: Average packet delivery ratio versus PU-OFF time $\lambda_{c,OFF}^{p}$ at α = 0.9 for TRL scheme in 6-node and 10-node topologies.

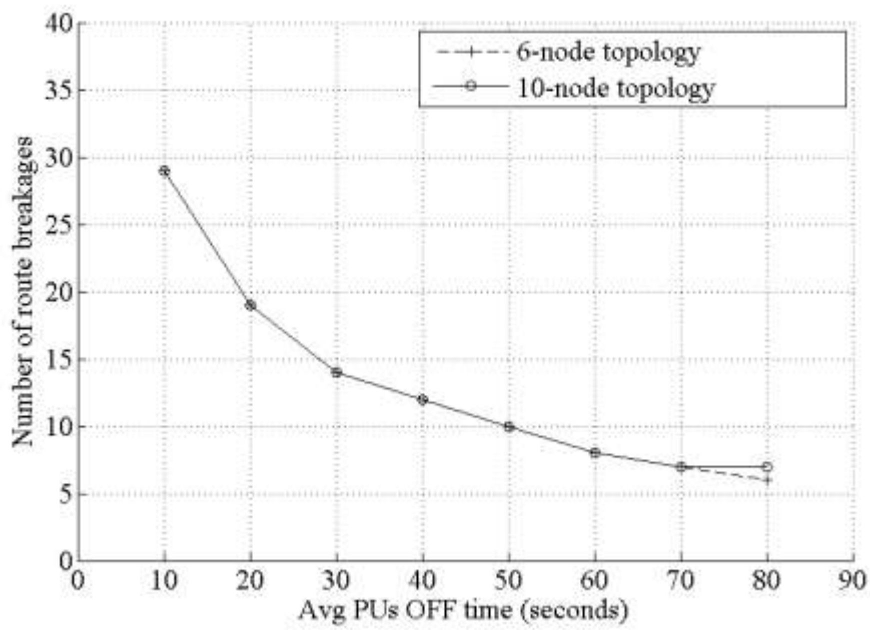


Figure 5-26: Average number of route breakages versus PU-OFF time $\lambda_{c,OFF}^{p}$ at α = 0.9 for TRL scheme in 6-node and 10-node topologies.

5.6 Chapter summary

This chapter proposed and implemented three route selection schemes on a USRP/ GNU Radio testbed environment based on reinforcement learning (RL) and spectrum leasing (SL) approaches in order to enhance the network performance of multi-hop cognitive radio (CR) networks. Specifically, the three schemes are the traditional RL (TRL) approach, the RL approach with average Q-value (ARL), and a SL approach. The RL-based schemes use an artificial intelligence technique, to make route selection; whereas, the SL-based scheme uses a spectrum occupancy map received from primary users (PUs) to make route selection. The RL-based schemes are best suited in situation when primary users (PUs) and secondary users (SUs) follow a distributed architectural model, in which each SUs has only the local spectrum knowledge. Whereas the SL-based scheme can be used when PUs and SUs follow a centralized architectural model, in which each SUs has only the local spectrum knowledge. Experimental results show that the proposed RL- and SL- based schemes select routes with the highest Q-value and highest minimum channel available time, respectively, contributing to lower number of route breakages and higher throughput and packet delivery ratio compared to a highest-channel (HC) scheme. The reason is that, in the RL-based schemes, intelligence is incorporated to find the estimated value of channel available time at the bottleneck link. Similarly, in the SL-based scheme, the SUs are aware of the PUs' activities and so the exact channel available time at the bottleneck link is used. Whereas, in the HC scheme, routes are selected on the basis of the highest number of available channels irrespective of the channel available time. The proposed RL-based and SL-based schemes appear to perform well as they provide enhanced performance in both 6-node and 10-node topologies. The results show that, 6-node topology achieves better performance than 10-node topology. Although, the 10-node topology has higher number of routes than 6-node topology, but the degradation in the performance is due to the higher number of hops in 10-node topology. This chapter also mathematically analyzed the complexity of route selection schemes, and found that the complexity of a multi-route with multi-hop networks increases with the increase in number of routes and number of hops, which may degrade the overall QoS performance (e.g., throughput and packet delivery ratio). The possibilities to extend this work in the future is discussed in chapter 6.

6.0 SUMMARY AND FUTURE WORK

6.1 Summary

Cognitive radio (CR) is the next-generation wireless communication system that has been proposed to address spectrum scarcity in the existing static assignment policy. CR uses dynamic spectrum access to solve this problem with two main approaches to access the channels. Firstly, SUs access the channel in an opportunistic manner using artificial intelligence so that SUs can learn and make decisions on channel and route selections in CR networks while the PUs are oblivious to the presence of the SUs. Secondly, SUs negotiate with PUs for channel access in a collaborative or partnership manner in order to achieve mutual benefit, such as Quality of Service (QoS) enhancement for both PUs and SUs.

This research work is a pioneering effort to examine opportunistic and collaborative channel access approaches at the network layer using real testbed implementation. There are three major contributions in this thesis. Firstly, a channel selection scheme is implemented in multi-hop CR network using reinforcement learning (RL) on a USRP/ GNU radio platform. Secondly, route selection schemes are implemented in a multi-hop CR network using RL and SL with the objective of improving QoS performance. Thirdly, addresses the challenges of network-layer implementation using USRP/ GNU radio platform. Analyzes the outcomes and results of the proposed schemes implemented on a USRP/ GNU radio platform. This thesis has achieved its overall goal. The four research questions given in section 1.4 are answered below:

What are the recent advances in spectrum leasing for CR networks and the real testbed implementation for the deployment of multi-hop communication in CR network?

Chapter 2 presents a comprehensive review on spectrum leasing schemes along with the limited work that has been implemented on a real testbed environment. The SL schemes are reviewed with an objective to determine the advantages, functionalities, characteristics and challenges of each scheme in CR networks. Spectrum leasing schemes have been shown to address the concerns poised to the traditional CR networks, so that PUs can enhance their network performance and maximize their monetary gain; while the SUs can enhance their network performance through exclusive access to white spaces. Examples of PU's gains are monetary gain and network performance enhancement; while example of SU's gain is dedicated channel access. To achieve these gains, PUs need to determine the cost of the white spaces, the PU's and SU's channel access time, SU's selection as a relay nodes, as well as PU's own packet transmission; while SUs need to select the appropriate PUs according to the SUs' QoS requirements and the cost of white spaces, as well as to determine channel access time between SUs. In the literature, the network topology of PUs and SUs can be either centralized or distributed; and the PUs and SUs operate among themselves using intra-cooperative and inter-cooperative modes, respectively. The challenges associated with PUs are the selection of the appropriate SUs to increase the monetary gain, the distribution of channel access time between PUs and SUs as well as continuous monitoring of SUs' activities; while the challenge associated to SUs is the selection of optimal channels in order to reap the benefits of spectrum leasing. Additionally, chapter 2 discusses various performance enhancement achieved by the spectrum leasing schemes (e.g. lower outage probability and higher outage capacity). Furthermore, open issues are recommended in order to spark new interests in this research area (e.g. enhancing auction and coordination mechanism and investigation of energy-efficient spectrum leasing schemes), as well as new kinds of CR networks such as CR sensor networks. Finally, at the end of chapter 2, the limited work implemented on the real testbed environment for multi-hop CRN has been discussed; specifically at the network layer.

How to design the system model that covers PUs' activities and SUs architecture, which is suitable and scalable for real testbed implementation?

Chapter 4 and chapter 5 discuss the architectures that have been proposed to model the activities of PUs and SUs in order to implement the concept of CRN using a real testbed environment.

Chapter 4 discusses the system model, which is comprises of two computers, four USRP/GNU units, and a gigabit Ethernet switch. All USRP units are connected through a switch to one of the computers. Three SUs and a single PU are configured through USRP which serve as the hardware platform. The two computers, namely the source computer and the destination computer use GNU Radio that serves as the software platform for USRPs. The SUs form a multi-hop network, and they are the source, relay and destination nodes, and transmit video packets from the source computer to the destination computer. The transmission from the source computer to the destination computer is accomplished through two wireless channels, specifically one channel for source to relay transmission and the other for relay to destination transmission, respectively. The PU transmits in any of the available channels.

Chapter 5 discusses the architecture of the USRP/ GNU radio platform for CR networks. There are two network architectures; one with six USRP SU nodes; while the other is with ten USRP SU nodes. The USRP provides the hardware RF frontend. GNU radio is a software interface written in Python, and it is installed in the computer. The GNU radio serves two main purposes. Firstly, it houses the decision making engine, and defines the operating environment (e.g., the PUs' activities, which are exponential ON/ OFF processes). Secondly, it provides a software interface to the USRP platform. The decision making engine is embedded in GNU radio software, and it is used to obtain decision making factors from the operating environment (e.g., PUs' activities), analyze the decision making factors, and make selection of the optimal action (e.g., route selection).

How to apply the reinforcement learning (RL) approach on a real testbed implementation?

Chapter 4 and chapter 5 further discuss the application of reinforcement learning (RL) to channel selection and route selection, respectively.

In chapter 4, a channel selection model is defined and implemented in a real testbed for the transmission between a source node and a destination node. As PU may appear in any of these channels, SUs need to select another channel for transmission. In this case, RL has been applied and implemented to tackle this channel selection challenge in a multi-hop CR network. The RL scheme is based on Q-model and is updated after every time window of 1 minute.

In chapter 5, three route selection schemes based on reinforcement learning (RL) and spectrum leasing (SL) are proposed and implemented on real testbed. The proposed RL-based route selection schemes are traditional RL scheme (TRL) and a RL scheme with average Q-value (ARL). In general, TRL and ARL share a similar algorithm except the way in which the Q-values are updated: TRL calculates the Q-values using the traditional approach, whereas ARL uses an average Q-value. Please note in the SL-based route selection scheme, the PUs collaborate with SUs by sharing their channel state information, hence it may not require to incorporate any artificial intelligence approach (e.g. RL).

How to improve the QoS performance of SUs in the proposed schemes using the centralized/ distributed models?

Chapter 4 and chapter 5 have shown that the proposed schemes improve the QoS performance of the SU networks.

Chapter 4 implemented RL-based channel selection mechanism on a multi-hop CR network that adopts the distributed model, for the selection of suitable channel with an objective of evading the PUs' activities in order to improve the throughput performance. The RL approach shown that whenever the transmission channel among the SUs is re-occupied with signals from PUs, the SUs update the Q-value of the available channels, and switch to another channel free from the PUs' activities; and thus achieves stable throughput performance.

Chapter 5 proposed and implemented three schemes that are based on RL and SL in 6-node and 10-node topologies, and the results are compared with Highest-Channel (HC) scheme. Overall, the SL-based scheme achieves higher throughput and packet delivery ratio, as well as lower number of route breakages, in comparison with the RL-based and HC schemes. This is because the SL-based scheme uses the centralized model, in which PUs share their respective spectrum occupancy map (i.e., $\lambda^{p}_{c_c,OFF}$) with the SUs located within their respective transmission range. Whereas, the RL schemes use the distributed model, and due to this reason they need to estimate the channel available time by themselves in order to compute the Q-value. In all cases, the HC scheme shows the least performance in comparison with SL- and RL-based schemes as it selects a route with the highest number of available channels, which may have low channel available time at its bottleneck link. Despite the fact that RL-based schemes use the distributed model, and are not aware about the PUs activities; however, the performance achieved by RL-based schemes are close to the performance of SL-based scheme. Furthermore, the QoS performance achieved by the 6-node and 10-node topologies are compared in order to investigate the scalability aspect of the schemes; and found that the proposed RL-based and SL-based schemes appear to scale well as they provide enhanced performance in both 6-node and 10-node topologies.

6.2 Future work

The section highlights the most significant directions for future work.

6.2.1 Considering the complete implementation of SL on real testbed

The open issues discussed in chapter 2 can be considered for the implementation of complete scenario of SL on a real testbed. For instance, the auction mechanism requires proper coordination in which the PUs (or SUs) make decisions on the selection of SUs (or PUs) participating in spectrum leasing so that both PUs and SUs mutually agree to fulfill each other requirements. This mechanism has been investigated using the simulation model and the mathematical model. However, no work has been done to investigate it on a testbed environment; so, it can be investigated on a real testbed implementation using a single-hop and a multi-hop CR network.

6.2.2 Suggested future work extensions of proposed route selection schemes (i.e., RL and SL)

The future plan is to extend the proposed work by relaxing the assumptions made in chapter 5, as well as to consider the typical phenomena (e.g., fading, shadowing, and distance among the nodes) in order to emulate a more real propagation environment. For instance, the extension of this work can be achieved by incorporating the channel sensing mechanism on a real testbed, which may degrade the QoS performance (e.g., throughput) of the proposed RL-based schemes, as well as HC scheme. Furthermore, there is a plan to increase the number of routes in order to better understand the computational performance of the different route selection schemes in real testbest, as well as to apply the multi-agent approach of RL so that every SU node is incorporated with intelligence to solve more complex scenarios.

REFERENCES

1. Devroye N, Vu M, and Tarokh V. Cognitive radio networks. IEEE Signal Processing Magazine 2008; 25: 12-23.
2. Hiertz GR, Denteneer D, Stibor PL, Zang Y, Costa XP, and Walke B. The IEEE 802.11 universe. IEEE Communications Magazine 2010; 48: 62-70.
3. Chan S, Vu HL, and Liu J. Performance analysis and optimization of best-effort service in IEEE 802.16 networks. Wiley Wireless Communications and Mobile Computing 2014; 14: 254-268.
4. Ferrus R, Sallent O, Baldini G, and Goratti L. LTE: the technology driver for future public safety communications. IEEE Communications Magazine 2-13; 51: 154-161.
5. Ghosh A, Ratasuk R, Mondal B, Mangalvedhe N, and Thomas T. LTE-advanced: next-generation wireless broadband technology [Invited Paper]. IEEE Wireless Communications 2010; 17: 10-22.
6. Akyildiz IF, Lee WY, Vuran MC, and Mohanty S. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey. Elsevier Computer Networks 2006; 50: 2127-2159.
7. Zhao Q, and Sadler BM. A survey of dynamic spectrum access. IEEE Signal Processing Magazine 2007; 24: 79-89.
8. Bolomey JC. Recent European developments in active microwave imaging for industrial, scientific, and medical applications. IEEE Transactions on Microwave Theory and Techniques 1989; 37: 2109-2117.
9. Jayaweera SK, Vazquez-Vilar G, and Mosquera C. Dynamic spectrum leasing: A new paradigm for spectrum sharing in cognitive radio networks. IEEE Transactions on Vehicular Technology 2010; 59: 2328-2339.
10. Oruc AY. Handbook of Scientific Proposal Writing. CRC Press 2011.
11. Kasch WT, Ward JR, and Andrusenko J. Wireless network modeling and simulation tools for designers and developers. IEEE Communications magazine 2009; 47: 120-127.
12. Rachedi A, Lohier S, Cherrier S, and Salhi I. Wireless network simulators relevance compared to a real testbed in outdoor and indoor environments. International Journal of Autonomous and Adaptive Communications Systems 2012; 5: 88-101.
13. Cacciapuoti AS, Caleffi M, and Paura L. Reactive routing for mobile cognitive radio ad hoc networks. Elsevier Ad Hoc Networks 2012; 10: 803-815.
14. Haykin S. Cognitive radio: brain-empowered wireless communications. IEEE Journal on Selected Areas in Communications 2005; 23: 201-220.
15. Liang YC, Chen KC, Li GY, and Mahonen P. Cognitive radio networking and communications: An overview. IEEE Transactions on Vehicular Technology 2011; 60: 3386-3407.
16. Chiarotto D, Simeone O, and Zorzi M. Spectrum leasing via cooperative opportunistic routing techniques. IEEE Transactions on Wireless Communications 2011; 10: 2960-2970.
17. Wei G, Xiang Y, Ji M, and Zhu P. An analytical model for optimal spectrum leasing under constraints of quality of service in CRNs. Elsevier Computer Networks 2014; 74: 71-80.

18. Chen Z, and Chen C. Adaptive energy-efficient spectrum probing in cognitive radio networks. Elsevier Ad Hoc Networks 2014; 13: 256-270.
19. Liang Q, Wang X, Tian X, Wu F, and Zhang Q. Two-dimensional route switching in cognitive radio networks: A game-theoretical framework. IEEE/ACM Transactions on Networking 2015; 23: 1053-1066.
20. Lopez-Benitez M, and Casadevall F. Time-dimension models of spectrum usage for the analysis, design, and simulation of cognitive radio networks. IEEE Transactions on Vehicular Technology 2013; 62: 2091-2104.
21. Atapattu S, Tellambura C, and Jiang H. Energy detection based cooperative spectrum sensing in cognitive radio networks. IEEE Transactions on Wireless Communications 2011; 10: 1232-1241.
22. El-Howayek G, and Jayaweera SK. Distributed dynamic spectrum leasing (D-DSL) for spectrum sharing over multiple primary channels. IEEE Transactions on Wireless Communications 2011; 10: 55-60.
23. Xue D, and Ekici E. Cross-layer scheduling for cooperative multi-hop cognitive radio networks. IEEE Journal on Selected Areas in Communications 2013; 31: 534-543.
24. Bourdena A, Mavromoustakis CX, Kormentzas G, Pallis E, Mastorakis G, and Yassein MB. A resource intensive traffic-aware scheme using energy-aware routing in cognitive radio networks. Future Generation Computer Systems 2014; 39: 16-28.
25. Backens J, Xin C, and Song M. A novel protocol for transparent and simultaneous spectrum access between the secondary user and the primary user in cognitive radio networks. Elsevier Computer Communications 2015; 69: 98-106.
26. Sun L, Zheng W, Rawat N, Sawant V, Koutsonnikolas D. Performance comparison of routing protocols for cognitive radio networks. IEEE Transaction on Mobile Computing 2014; 14: 1272-1286.
27. Huang X, Lu D, Li P, Fang Y. Coolest path: spectrum mobility aware routing metrics in cognitive ad hoc networks. In: Proceedings of 31st International Conference on Distributed Computing Systems (ICDCS). IEEE; 2011. p. 182–191.
28. Nagaraju PB, Ding L, Melodia T, Batalama SN, Pados DA, and Matyjas JD. Implementation of a distributed joint routing and dynamic spectrum allocation algorithm on USRP2 radios. In: Proceedings of 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), IEEE; p. 1-2.
29. Zhou P, Chang Y, and Copeland J. Reinforcement learning for repeated power control game in cognitive radio networks. IEEE Journal on Selected Areas in Communications 2012; 30: 54-69.
30. Yau KLA, Komisarczuk P, and Teal PD. Enhancing network performance in distributed cognitive radio networks using single-agent and multi-agent reinforcement learning. In: Proceedings of IEEE 35th Conference on Local Computer Networks (LCN). IEEE; 2010. p. 152-159.
31. Teng Y, Zhang Y, Niu F, Dai C, and Song M. Reinforcement learning based auction algorithm for dynamic spectrum access in cognitive radio networks. In: Proceedings of IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall), IEEE; 2010. p. 1-5.
32. Wang J, Ghosh M, and Challapali K. Emerging cognitive radio applications: A survey. IEEE Communications Magazine 2011; 49: 74-81.

33. Syed AR, and Yau KLA. On cognitive radio-based wireless body area networks for medical applications. In: Proceedings of IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE). IEEE; 2013. p. 51-57.
34. Shao C, Roh, H, and Lee W. Aspiration level-based strategy dynamics on the coexistence of spectrum cooperation and leasing. IEEE Communications Letters 2014; 18: 70-73.
35. Chen G, Xiang Z, Xu C, and Tao M. On degrees of freedom of cognitive networks with user cooperation. IEEE Wireless Communications Letters 2012; 1: 617-620.
36. Wu Y, Wang B, Liu KR, and Clancy TC. Collusion-resistant multi-winner spectrum auction for cognitive radio networks. In: Proceedings of IEEE GLOBECOM Global Telecommunications Conference, IEEE; 2008. p. 1-5.
37. Jayaweera SK, Bkassiny M, and Avery K. Asymmetric cooperative communications based spectrum leasing via auction in cognitive radio networks," IEEE Transactions on Wireless Communications 2011; 10: 2716-2724.
38. Simeone O, Gambini J, Bar-Ness Y, and Spagnolini U. Cooperation and cognitive radio. In: Proceedings of IEEE International Conference on Communications, IEEE; 2007. p. 6511-6515.
39. Cao B, Zhang Q, Mark JW, Cai LX, and Poor HV. Toward efficient radio spectrum utilization: user cooperation in cognitive radio networking. IEEE Network 2012; 26: 46-52.
40. Oey CH, Christian I, and Moh S. Energy-and cognitive-radio-aware routing in cognitive radio sensor networks. International Journal of Distributed Sensor Networks 2012; 1-11.
41. Syed AR, and Yau KLA. Spectrum leasing in cognitive radio networks: A survey. International Journal of Distributed Sensor Networks 2014; 2014: 1-22.
42. Yi Y, Zhang J, Zhang Q, Jiang T, and Zhang J. Cooperative communication-aware spectrum leasing in cognitive radio networks. In: Proceedings of IEEE Symposium on New Frontiers in Dynamic Spectrum, IEEE; 2010. p. 1-11.
43. Zhou H, Liu B, Gui L, Wang X, and Li Y. Fast spectrum sharing for cognitive radio networks: a joint time-spectrum perspective. In: Proceedings of IEEE GLOBECOM Global Telecommunications Conference, IEEE; 2011. p. 1-5.
44. Kasbekar GS, and Sarkar S. Spectrum pricing games with spatial reuse in cognitive radio networks. IEEE Journal on Selected Areas in Communications 2012; 30: 153-164.
45. Song H, and Lin X. A leasing oriented MAC protocol for high spectrum usage in cognitive radio networks. In: Proceedings of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, IEEE; 2009. p. 173-178.
46. Cheng SM, Lin P, Huang DW, and Yang SR. A study on distributed/centralized scheduling for wireless mesh network. In: Proceedings of the 2006 international conference on Wireless communications and mobile computing, ACM; 2006. p. 599-604.
47. Sodagari S, Attar A, and Bilén SG. On a truthful mechanism for expiring spectrum sharing in cognitive radio networks. IEEE Journal on Selected Areas in Communications 2011; 29: 856-865.

48. Alptekin GI, and Bener AB. An efficient spectrum management mechanism for cognitive radio networks. In: Proceedings of IEEE International Symposium on Integrated Network Management, IEEE; 2009. p. 653-660.
49. Chan YW, Chang RY, Chien FT, Chang MK, and Chung YC. Coalitional game formulation for multi-channel cooperative cognitive radio networks. In: Proceedings of IEEE GLOBECOM Workshops (GC Wkshps), IEEE; 2011. p. 975-980.
50. Huang SW, Chan YW, Orien FT, and Chung YC. Efficient resource allocation in cooperative cognitive radio networks: A coalitional game approach. In: Proceedings of IET International Conference on Wireless Mobile and Computing, IEEE; 2011. p. 133-138.
51. Jayaweera SK, and Li T. Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games. IEEE Transactions on Wireless Communications 2009; 8: 3300-3310.
52. Vazquez-Vilar G, Mosquera C, and Jayaweera SK. Primary user enters the game: Performance of dynamic spectrum leasing in cognitive radio networks. IEEE Transactions on Wireless Communications 2010; 9: 1-5.
53. Nie PY, and Zhang PA. A note on Stackelberg games. In: Proceedings of IEEE Chinese Control and Decision Conference, IEEE; 2008. p. 1201-1203.
54. Scutari G, and Pang JS. Joint sensing and power allocation in nonconvex cognitive radio games: Nash equilibria and distributed algorithms. IEEE Transactions on Information Theory 2013; 59: 4626-4661.
55. Yau KLA, Komisarczuk P, and Teal PD. Reinforcement learning for context awareness and intelligence in wireless networks: Review, new features and open issues. Journal of Network and Computer Applications 2012; 35: 253-267.
56. Le TH, and Van Barel M. A convex optimization method to solve a filter design problem. Journal of Computational and Applied Mathematics 2014; 255: 183-192.
57. Lin S, and Fang X. Two-level game based spectrum lease framework in cognitive radio networks. In: Proceedings of 11th IEEE Singapore International Conference on Communication Systems, IEEE; 2008. p. 104-108.
58. Kim H, and Shin KG. Optimal admission and eviction control of secondary users at cognitive radio hotspots. In: Proceedings of 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, IEEE; 2009. p. 1-9.
59. Murawski R, and Ekici E. Utilizing dynamic spectrum leasing for cognitive radios in 802.11-based wireless networks. Elsevier Computer Networks 2011; 55: 2646-2657.
60. Toroujeni SMM, Sadough SMS, and Ghorashi SA. On time-frequency resource leasing in cognitive radio networks. Wiley Wireless Personal Communications 2012; 65: 583-600.
61. Chen PY, Ao WC, Lin SC, and Chen KC. Reciprocal spectrum sharing game and mechanism in cellular systems with cognitive radio users. In: Proceedings of IEEE GLOBECOM Workshops (GC Wkshps), IEEE; 2011. p. 981-985.
62. Lei WANG, Xu XK, Xu WJ, He ZQ, and Lin JR. Novel pricing model for spectrum leasing in secondary spectrum market. The Journal of China Universities of Posts and Telecommunications 2010; 17: 12-19.

63. Stanojev I, Simeone O, Bar-Ness Y, and Yu T. Spectrum leasing via distributed cooperation in cognitive radio. In: Proceedings of IEEE International Conference on Communications, IEEE; 2008. p. 3427-3431.
64. Wang H, Gao L, Gan X, Wang X, and Hossain E. Cooperative spectrum sharing in cognitive radio networks: A game-theoretic approach. In: Proceedings of IEEE International Conference on Communications, IEEE; 2010. p. 1-5.
65. Zhang B, Chen K, Gou X, and Cheng S. Spectrum leasing via selective cooperation in distributed cognitive radio networks. In: Proceedings of 2nd IEEE International Conference on Network Infrastructure and Digital Content, IEEE; 2010. p. 16-20.
66. Zhu K, Niyato D, Wang P, and Han Z. Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks. IEEE Transactions on Wireless Communications 2012; 11: 1136-1145.
67. Asadumzzaman, Kong HY, and Koo I. Opportunistic relaying based spectrum leasing for cognitive radio networks. Journal of Communications and Networks 2011; 13: 50-55.
68. Khalil K, Karaca M, Ercetin O, and Ekici E. Optimal scheduling in cooperate-to-join cognitive radio networks. In: Proceedings of IEEE INFOCOM, IEEE; 2011. p. 3002-3010.
69. Hakim K, Jayaweera SK, El-Howayek G, and Mosquera C. Efficient dynamic spectrum sharing in cognitive radio networks: Centralized dynamic spectrum leasing (C-DSL). IEEE Transactions on Wireless Communications 2010; 9: 2956-2967.
70. Li T, Tang T, and Chang C. A new backoff algorithm for IEEE 802.11 distributed coordination function. In: Proceedings of Sixth International Conference on Fuzzy Systems and Knowledge Discovery, IEEE; 2009. p. 455-459.
71. Barros DJ, Wilson SK, and Kahn JM. Comparison of orthogonal frequency-division multiplexing and pulse-amplitude modulation in indoor optical wireless links. IEEE Transactions on Communications 2012; 60: 153-163.
72. Kvaternik K, and Pavel L. Lyapunov analysis of a distributed optimization scheme. In: Proccedings of 5th International Conference on Network Games, Control and Optimization (NetGCooP), IEEE; 2011. p. 1-5.
73. Liu Y, Xu D, Feng Z, and Zhang P. Outage capacity of cognitive radio in Rayleigh fading environments with imperfect channel information. Journal of Information and Computational Science 2012; 9: 955-968.
74. Yan S, Peng M, Wang W, Dong L, and Ahmed M. Relay self-selection for secure cooperative in Amplify-And-Forward networks. In: Proceedings of 7th International ICST Conference on Communications and Networking in China (CHINACOM), IEEE; 2012. p. 581-585.
75. Cardoso LS, Massouri A, Guillon B, Hutu F, Villemaud G, Risset T, and Gorce JM. CorteXlab: A Cognitive Radio Testbed for Reproducible Experiments. In: Proceedings of Wireless @ Virginia Tech Symposium, 2014. p. 1-7.
76. Le Nir V, and Scheers B. Implementation of an adaptive OFDMA PHY/MAC on USRP platforms for a cognitive tactical radio network. In: Proceedings of Military Communications and Information Systems Conference (MCC), IEEE; 2012. p. 1-7.
77. Yan Z, Ma Z, Cao H, Li G, and Wang W. Spectrum sensing, access and coexistence testbed for cognitive radio using USRP. In: Proceedings of 4th International

Conference on Circuits and Systems for Communications (ICCSC), IEEE; p. 270-274.
78. Yang L, Zhang Z, Hou W, Zhao BY, and Zheng H. Papyrus: a software platform for distributed dynamic spectrum sharing using SDRs. ACM SIGCOMM Computer Communication Review 2011; 41: 31-37.
79. Zhang Q, Jia J, and Zhang J. Cooperative relay to improve diversity in cognitive radio networks. IEEE Communications Magazine 2009; 47: 111-117.
80. Cesana M, Cuomo F, and Ekici E. Routing in cognitive radio networks: Challenges and solutions. Elsevier Ad Hoc Networks 2011; 9: 228-248.
81. Bayhan S, and Alagoz F. Scheduling in centralized cognitive radio networks for energy efficiency. IEEE Transactions on Vehicular Technology 2013; 62: 582-595.
82. Kordali AV, and Cottis PG. A Contract-Based Spectrum Trading Scheme for Cognitive Radio Networks Enabling Hybrid Access. IEEE Access 2015; 3: 1531-1540.
83. Di Felice M, Chowdhury KR, Kim W, Kassler A, and Bononi L. End-to-end protocols for cognitive radio ad hoc networks: An evaluation study. Performance Evaluation 2011; 68: 859-875.
84. Patil DP, and Wadhai VM. NS2 based advanced routing model for cognitive radio networks from dynamic spectrum management perception. In: Proceedings of IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS), IEEE; p. 1-5.
85. Qadir J. Artificial intelligence based cognitive routing for cognitive radio networks. Artificial Intelligence Review 2016; 45: 25-96.
86. Pefkianakis I, Wong SH, and Lu S. SAMER: Spectrum aware mesh routing in cognitive radio networks. In: Proceedings of 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, IEEE; 2008. p. 1-5.
87. How KC, Ma M, and Qin Y. Routing and QoS provisioning in cognitive radio networks. Elsevier Computer Networks 2011; 55: 330-342.
88. Chowdhury KR, and Akyildiz IF. CRP: A routing protocol for cognitive radio ad hoc networks. IEEE Journal on Selected Areas in Communications 2011; 29: 794-804.
89. Deng S, Chen J, He H, and Tang W. Collaborative strategy for route and spectrum selection in cognitive radio networks. In: Proceedings of Future generation communication and networking (FGCN), IEEE; 2007. p. 168-172.
90. Zhu GM, Akyildiz IF, and Kuo GS. STOD-RP: A spectrum-tree based on-demand routing protocol for multi-hop cognitive radio networks. In: Proceedings of Global Telecommunications Conference (GLOBECOM), IEEE; 2008. p. 1-5.
91. Bogale TE, and Vandendorpe L. USRP implementation of Max-Min SNR signal energy based spectrum sensing algorithms for cognitive radio networks. In: Proceedings of IEEE International Conference on Communications (ICC), IEEE; 2014. p. 1478-1482.
92. Zhao YC, Wu J, Li WZ, and Lu SL. Throughput optimization in cognitive radio networks ensembling physical layer measurement. Journal of Computer Science and Technology 2015; 30: 1290-1305.
93. Reyes H, Subramaniam S, Kaabouch N, and Hu WC. A spectrum sensing technique based on autocorrelation and euclidean distance and its comparison with energy

detection for cognitive radio networks. Computers & Electrical Engineering 2016; 52: 319-327.

94. Qi J, Hu F, Li X, Koushik AM, Hu L, and Kumar S. CR Based Video Communication Testbed with Robust Spectrum Sensing/Handoff. In: Proceedings of Information Technology: New Generations. Advances in Intelligent Systems and Computing, Springer International Publishing; 2016. p. 59-70
95. Reddy DN, Ravinder Y, and Sharma C. Testbed implementation of multi dimensional spectrum sensing schemes for cognitive radio. ICTACT Journal on Communication Technology 2016; 7: 1289-1294.
96. Wawrzyński P. Learning to control a 6-degree-of-freedom walking robot. In: Proceedings of the International Conference on Computer as a Tool EUROCON, IEEE; 2007. p. 698-705.
97. Kaelbling LP, Littman ML, and Moore AW. Reinforcement learning: A survey. Journal of artificial intelligence research 1996; 237-285.
98. Boushaba M, Hafid A, and Belbekkouche A. Reinforcement learning-based best path to best gateway scheme for wireless mesh networks. In: Proceedings of 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), IEEE; 2011. p. 373-379.
99. Felice MD, Chowdhury KR, Kassler A, and Bononi L. Adaptive sensing scheduling and spectrum selection in cognitive wireless mesh networks. In: Proceedings of 20th international conference on Computer communications and networks (ICCCN), IEEE; 2011. p. 1-6.
100. Li Y, Ji H, Li X, and Leung V. Dynamic channel selection with reinforcement learning for cognitive WLAN over fiber. Wiley International Journal of Communication Systems 2012; 25: 1077-1090.
101. Galindo-Serrano A, and Giupponi L. Distributed Q-learning for aggregated interference control in cognitive radio networks. IEEE Transactions on Vehicular Technology 2010; 59: 1823-1834.
102. Wu C, Chowdhury K, Di Felice M, and Meleis W. Spectrum management of cognitive radio using multi-agent reinforcement learning. In: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Industry track, 2010. p. 1705-1712.
103. Jiang T, Grace D, and Mitchell PD. Efficient exploration in reinforcement learning-based cognitive radio spectrum sharing. IET communications 2011; 5: 1309-1317.
104. Simeone O, Stanojev I, Savazzi S, Bar-Ness Y, Spagnolini U, and Pickholtz R. Spectrum leasing to cooperating secondary ad hoc networks. IEEE Journal on Selected Areas in Communications 2008; 26: 203-213.
105. Vazquez-Vilar G, Mosquera C, and Jayaweera SK. Primary user enters the game: Performance of dynamic spectrum leasing in cognitive radio networks. IEEE Transactions on Wireless Communications 2010; 9: 1-5.
106. Golrezaei N, Molisch, AF, Dimakis AG, and Caire G. Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution. IEEE Communications Magazine 2013; 51: 142-149.
107. Jiang Q, Leung VC, Pourazad MT, Tang H, and Xi HS. Energy-efficient adaptive transmission of scalable video streaming in cognitive radio communications. IEEE Systems Journal 2016; 10: 761-772.

108. Saki H, and Shikh-Bahaei M. Cross-layer resource allocation for video streaming over OFDMA cognitive radio networks. IEEE Transactions on Multimedia 2015; 17: 333-345.
109. Saki H, Martini MG, and Shikh-Bahaei M. Multi-user scalable video transmission over cognitive radio networks. In: Proceedings of IEEE International Conference on Communications (ICC), IEEE; 2015. p. 7564-7569.
110. Shah GA, Alagoz F, Fadel EA, and Akan OB. A spectrum-aware clustering for efficient multimedia routing in cognitive radio sensor networks. IEEE Transactions on Vehicular Technology 2014; 63: 3369-3380.
111. Huang Z, Wang W, and Zhang Y. Design and implementation of cognitive radio hardware platform based on USRP. In: Proceedings of IET International Conference on Communication Technology and Application, IEEE; 2011. p. 160-164.
112. Gavrilovska L, and Atanasovski V. Spectrum sensing framework for cognitive radio networks. Wiley Wireless Personal Communications 2011; 59: 447-469.
113. Popoola JJ, and Van Olst R. The performance evaluation of a spectrum sensing implementation using an automatic modulation classification detection method with a Universal Software Radio Peripheral. Elsevier Expert Systems with Applications 2013; 40: 2165-2173.
114. Chen Z, Guo N, Hu Z, and Qiu RC. Experimental validation of channel state prediction considering delays in practical cognitive radio. IEEE Transactions on Vehicular Technology 2011; 60: 1314-1325.
115. Chantaraskul S, and Moessner K. Implementation of a genetic algorithm-based decision making framework for opportunistic radio. IET communications 2010; 4: 495-506.
116. Kryszkiewicz P, Bogucka H, and Wyglinski AM. Protection of primary users in dynamically varying radio environment: practical solutions and challenges. EURASIP Journal on wireless communications and networking 2012; 2012: 1-20.
117. Tahir M, Mohamad H, Ramli N, and Jarot SP. Experimental implementation of dynamic spectrum access for video transmission using USRP. In: Proceedings of International Conference on Computer and Communication Engineering (ICCCE), IEEE; 2012. p. 228-233.
118. Ahn C, Bang S, Kim H, Lee S, Kim J, Choi S, and Glossner J. Implementation of an SDR system using an MPI-based GPU cluster for WiMAX and LTE. Springer Analog Integrated Circuits and Signal Processing 2012; 73: 569-582.
119. Zhou R, Mian O, Li X, Wang B, and Wu Z. A software-defined radio based cognitive radio demonstration over FM band. Wiley Wireless Communications and Mobile Computing 2010; 10: 4-15.
120. Puschmann A, Kalil MA, and Mitschele-Thiel A. Implementation and evaluation of a practical SDR testbed. In: Proceedings of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management, ACM; p. 1-5.
121. Huang Y, Walsh PA, Li Y, and Mao S. A distributed polling service-based MAC protocol testbed. International Journal of Communication Systems 2014; 27: 3901-3921.
122. El-Hajjar M, Nguyen QA, Maunder RG, and Ng SX. Demonstrating the practical challenges of wireless communications using USRP. IEEE Communications Magazine 2014; 52: 194-201.

123. Briand A, Albert BB, and Gurjao EC. Complete software defined RFID system using GNU radio. In: Proceedings of International Conference on RFID-Technologies and Applications (RFID-TA), IEEE; 2012. p. 287-291.
124. Rovcanin M, De Poorter E, Moerman I, and Demeester P. A reinforcement learning based solution for cognitive network cooperation between co-located, heterogeneous wireless sensor networks. Elsevier Ad Hoc Networks 2014; 17: 98-113.
125. Coutinho N, Matos R, Marques C, Reis A, Sargento S, Chakareski J, and Kassler A. Dynamic dual-reinforcement-learning routing strategies for quality of experience-aware wireless mesh networking. Elsevier Computer Networks 2015; 88: 269-285.
126. Danieletto M, Quer G, Rao RR, and Zorzi M. CARMEN: a cognitive networking testbed on Android OS devices. IEEE Communications Magazine 2014; 52: 98-107.
127. Truong NB, Suh YJ, and Yu C. Latency analysis in GNU radio/USRP-based software radio platforms. In: Proceedings of IEEE Military Communications Conference (MILCOM), IEEE; 2013. p. 305-310.
128. Chen R, Park JM, and Bian K. Robust distributed spectrum sensing in cognitive radio networks. In: Proceedings of 27th Conference on Computer Communications (INFOCOM), IEEE; 2008.
129. Wang X, Sheng M, Zhai D, Li J, Mao G, and Zhang Y. Achieving Bi-Channel-Connectivity with Topology Control in Cognitive Radio Networks. IEEE Journal on Selected Areas in Communications 2014; 32: 2163-2176.
130. Sun H, Nallanathan A, Cui S, and Wang CX. Cooperative wideband spectrum sensing over fading channels. IEEE Transactions on Vehicular Technology 2016; 65: 1382-1394.
131. Atawi IE, Badarneh OS, Aloqlah MS, and Mesleh R. Energy-detection based spectrum-sensing in cognitive radio networks over multipath/shadowed fading channels. In: Proceedings of Wireless Telecommunications Symposium (WTS), IEEE; 2015. p. 1-6.
132. Rehmani MH, Viana AC, Khalife H, and Fdida S. Surf: A distributed channel selection strategy for data dissemination in multi-hop cognitive radio networks. Elsevier Computer Communications 2013; 36: 1172-1185.
133. Boyan JA, and Littman ML. Packet routing in dynamically changing networks: A reinforcement learning approach. Advances in neural information processing systems 1994; 671-671.
134. Al-Rawi HA, Yau KLA, Mohamad H, Ramli N, and Hashim W. Reinforcement Learning for Routing in Cognitive Radio Ad Hoc Networks. The Scientific World Journal 2014, 2014.
135. Arroyo-Valles R, Alaiz-Rodriguez R, Guerrero-Curieses A, and Cid-Sueiro J. Q-probabilistic routing in wireless sensor networks. In: Proceedings of 3rd International Conference on Intelligent Sensors, Sensor Networks and Information (ISSNIP), IEEE; 2007. p. 1-6.
136. VERT900 Antenna, http://www.ettus.com/product/details/VERT900 2015.
137. Xia B, Wahab MH, Yang Y, Fan Z, and Sooriyabandara M. Reinforcement learning based spectrum-aware routing in multi-hop cognitive radio networks. In: Proceedings of 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM). Hannover, Germany: IEEE; 2009. p. 1-5.