2022

# Efficient channel allocation schemes for multi-band relay networks

Mughal, Bushra

# Efficient Channel Allocation Schemes for Multi-band Relay Networks

by

Bushra Mughal

Thesis presented to the faculty of Electrical and Computer Engineering Department,

Lakehead University, Thunder Bay, ON, Canada,

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Ph.D.) in Electrical and Computer Engineering.

Lakehead University

Thunder Bay, ON, Canada.

# EXAMINING COMMITTEE MEMBERSHIP

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

| | |
|---|---|
| Supervisor: | Dr. Salama Ikki<br>Associate Professor, Dept. of Electrical Engineering,<br>Lakehead University |
| Co-Supervisor: | Dr. Zubair Fadlullah<br>Associate Professor, Dept. of Computer Science,<br>Lakehead University |
| Co-Supervisor: | Dr. Mostafa Fouda<br>Associate Professor, Dept. of Electrical and Computer Engineering,<br>Idaho State University, USA |
| Internal Member: | Dr. Waleed Ejaz<br>Associate Professor, Dept. of Electrical Engineering,<br>Lakehead University |
| Internal Member: | Dr. Thangarajah Akilan<br>Assistant Professor, Dept. of Electrical Engineering,<br>Lakehead University |
| External Member: | Dr. Khaled Rabie<br>Senior Lecturer, Dept. of Engineering,<br>Manchester Metropolitan University, UK |

## AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## ABSTRACT

As next generation wireless and mobile networks continue to evolve, the scarcity of frequency spectrum and bandwidth are on the rise. While ultra-dense small-scale networks aim to satisfy the demand for ultra-high data rate needs, they offer limited coverage to mobile users. While relay-based network coverage expansion emerges as an appealing solution, it typically contributes to the additional transmission delay between a source node (SN) and a destination node (DN). To support high data rate communication and reduce the transmission delay, the heterogeneous frequency bands could be adaptively/simultaneously used to transmit the data packets over the relay node/s (RN/s). In this thesis, the exploitation of heterogeneous frequency bands in this manner is referred to as the *multi-band communication*.

In this thesis, the advantages and shortcomings of relay-based transmission with multi-band communication are studied in a systematic manner. By first formulating the theoretical problem, it is then investigated that how to develop efficient approaches and algorithms to effectively reduce the packet relay latency. In particular, the primary concern is the added delay due to relaying data packets from SN till DN via one or more RNs and any additional coordination among nodes might end up in additional computational overhead. Therefore, there emerges a trade-off between optimal (i.e., best) and fast decisions. To address the trade-off between the quality and speed of the required solution, both centralized and distributed methods are analyzed and proposed, so as to reduce the packet latency from SN to DN. In centralized optimization approach, as the number of RNs increase, an optimal solution is difficult to determine because of non-deterministic computation time. On the other hand, if the distributed approach is used, the data communication latency can

be significantly reduced with the use of local decisions that are feasible and fast. Moreover, the concept of cognitive radios and spectrum sensing is also considered in this work to utilize multi-band communication on freely sensed available channels instead of over-utilizing the scarce spectrum.

First a heuristic-based, distributed channel allocation approach is proposed as a simple, baseline method to quickly compute acceptable solutions to reduce data packet latency for a multi-hop (relay) network exploiting multi-band communication. In this proposed greedy algorithm, each RN is able to make localized decisions regarding the best channel allocation over heterogeneous bands. Its adaptability analysis and computer-based simulations demonstrate that the proposed greedy approach outperforms the centralized approach and several other baseline (conventional) methods in terms of reduced relay latency.

It is then further investigated that, the initially proposed distributed heuristic can be used as a reference since it provides the first intuitive and distributed solution to the formulated problem. However, due to its greedy nature, there is scope for further improvement with regards to the quality of solutions. Since the greedy approach does not offer the performance bound guarantee, a better online distributed band/channel allocation strategy is explored by proposing a sequential algorithm based on game theory. Based on extensive computer-based simulation results, it is demonstrated that the proposed game-theoretic approach significantly outperforms the bipartite-based centralized oracle and other traditional methods in terms of a significantly lower packet latency with a derivable performance guarantee.

As an extension to the work done on game theory, its corresponding model-specific computational and theoretical analysis is also provided in this thesis. Game theoretic based models are always analyzed and justified with an important concept of Nash Equilibrium. For sequential game algorithms, Nash equilibrium is obtained by breaking down the game in sub-games, with two players each, and finding optimal solutions in a backward manner. This phenomena is known as sub-game perfect equilibrium using backward induction. The analysis of proposed model is presented with a simple yet illustrative example, based on which two numerical/tabular examples are also provided for picturing the practical side of this phenomena. It is then deduced that the sequential game can be converged and stabilized by obtaining the Nash Equilibria of the sub-games within the main game.

While existing greedy heuristics and game-theoretic techniques, developed for multi-band channel allocation, achieved minimum packet latency, however their performance drops significantly when network dynamism is introduced in terms of user mobility and non-quasi-static channel conditions. To handle such situations, customized machine learning algorithm is proposed so as to make RNs smarter (rather being only efficient) for making decisions. The problem is modeled as a distributed Markov Decision Process (MDP) involving various learning states for each RN. Since solving a MDP, traditionally consumes much time and is intractable for the RNs, the problem is reformulated as a reinforcement learning-based, smart channel adaptation problem, which is then optimally solved by a customized Q-Learning algorithm geared with $\epsilon$-greedy policy. Extensive computer-based simulation results demonstrate that proposed reinforcement learning algorithm outperforms the existing methods in terms of transmission time, buffer overflow, and effective throughput. The detailed convergence analysis of the proposed algorithm is also provided, by systematically finding and setting the appropriate parameters.

## ACKNOWLEDGEMENTS

First and foremost, I thank Allah (SWT) for blessing me with completion of this thesis and my degree.

No words are going to be enough to express my gratitude for all the Dua's of my father, Wing Commander Muhammad Nawaz Mughal (Ret.) and my mother, Razia Sultana. Nothing in life could have been possible without your endless trust and encouragement.

I would like to express my sincere gratitude to my supervisor, Dr. Salama Ikki, and my co-supervisors, Dr. Zubair Fadlullah and Dr. Mostafa Fouda, for putting their trust in me and for accepting to supervise this work as well as for their continued support, invaluable advice and encouragement throughout my PhD period. This work would have not been brought to completion without their expertise in formulating the research problems and methodology. Their insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I also extend my thanks to the committee members, for agreeing to judge this work as reviewers and to give constructive remarks that allowed me to improve this thesis.

## <u>DEDICATION</u>

I hereby dedicate my thesis and all my accomplishments to my parents who mean the world to me:

***Wing Commander Muhammad Nawaz Mughal (Ret.)***

***and***

***Mrs. Razia Sultana***

# Table of Contents

# List of Figures

# List of Tables

# Acronyms

**AWGN** Additive White Gaussian Noise 91

**BOC** Buffer Overflow Condition 36, 39–41

**CA** Congestion-Aware 17, 18, 77, 81–83

**CN** Central Node 9, 16, 52, 56, 57

**CR** Cognitive Radio ix, 4, 9

**CRN** Cognitive Radio Network 19, 20

**CSS** Cooperative Sensing Scheduling 20

**D2D** Device 2 Device 25, 122

**DF** Decode and Forward 122

**ISM** Industrial, Scientific and Medical 3

**MDP** Markov Decision Process ix, xii, xiii, xvii, 8, 20, 89, 90, 96, 97, 99–102, 104, 106, 107, 109, 120

**mmWave** milli-meter Wave 3, 20

**PACA** Power-Aware Congestion-Aware 18, 77, 81–83

**QoS** Quality of Service 2, 3, 5, 22, 28, 56, 62, 65, 96, 102, 104

**SINR** Signal to Interference Noise Ratio 17

**SNR** Signal to Noise Ratio ix, 2, 9, 10, 17, 21, 22, 30, 90, 91, 95, 101–104, 108, 111

**VLC** Visible Light Communication 3

**WLAN** Wireless Local Area Network 16

# Chapter 1

# Introduction and Problem Statement

In this era of high speed data communications and increasing number of users, there is a need to efficiently utilize the frequency spectrum along with increased user coverage. This can be obtained by utilizing advanced hardware radios, capable of multi-band/channel communication. Such advanced devices, when used in a relay network approach, result in increasing the coverage and reducing data latency drastically. This thesis provides detailed analysis of multi-band capable relay network and presents multiple algorithms showing the reduction in latency and its affect on effective throughput of the system.

This chapter comprises of the introduction to the terms, techniques, ideas and approaches which build the foundation of this thesis. It also provides details on the principal assumptions used throughout the work, which then leads towards the main objective and breakdown of the proposed thesis.

## 1.1 Introduction

This section provides basic definitions to the core ideas of this thesis: Relay network topology, multi-band communication, multi-band packet relaying, considered network approaches, graph theory.

### 1.1.1 Relay Network Communication

A very well-known broad class of network topology is being commonly used and researched in wireless networks: the relay network topology [1]. In this topology, the sender node (SN) sends data to a relay node (RN), which then re-transmits the same data to its destination node (DN). In a similar way, the SN could have more than one RN, relaying the same data packet. In effect, data relaying is used in two main scenarios:

1. *Increasing Transmission Range.* In a case where either the SN or DN are out of each other's transmission range, the packet relaying can be used to increase the transmission range. In this case, SN sends data to its chosen RN and by using either single or multiple RNs in between, the SN can forward the data packet to its DN, hence now increasing the transmission range.

2. *Worst Channel Conditions.* In a case where SN and DN are in transmission range but the direct channel between the sender and the destination is not clear enough (to achieve the minimum required Signal to Noise Ratio (SNR) to maintain the desired Quality of Service (QoS), SN can opt to relay data packet to DN using single or multiple RNs.

Relay network topology is effectively used to increase the coverage of the network and to match with minimum QoS requirements, however with the unfortunate cost of a delay in the total transmission time from the SN to DN. To mitigate this drawback, multi-band communications is utilized as discussed below.

### 1.1.2 Multi-band Communication

With advanced wireless mobile devices and systems, the world now have multi-band devices that can support a wide range of wireless communication standards and in turn a wide range of frequency bands that behave differently to path loss, fading, mobile blocking and other physical phenomena. Examples include milli-meter Wave (mmWave), Visible Light Communication (VLC), IEEE 802.11ax, marketed as Wi-Fi 6 by Wi-Fi Alliance [2]. IEEE 802.11ax is designed to operate in all Industrial, Scientific and Medical (ISM) bands between 1 and 6 GHz, in addition to the 2.4 and 5 GHz bands already allocated. Multiband communications not only helps in making mobile devices compatible with different wireless standards but can also be utilized to reduce the data transmission delay between two nodes as explained in next section.

### 1.1.3 Multi-band Relay Communication

Conventionally, in relay communication, if the SN transmits packets to the first RN, denoted as $\mathrm{RN}_1$ in Fig. 1.1, at $i^{\text{th}}$ channel of the $j^{\text{th}}$ band ($\mathrm{ch}_{sr_1}^{ij}$), $\mathrm{RN}_1$ waits for the same channel to become available (for re-transmitting packet to DN) as SN finishes sending its data packets completely. Here $s$ and $r_1$ refer to SN and $\mathrm{RN}_1$, respectively for brevity of

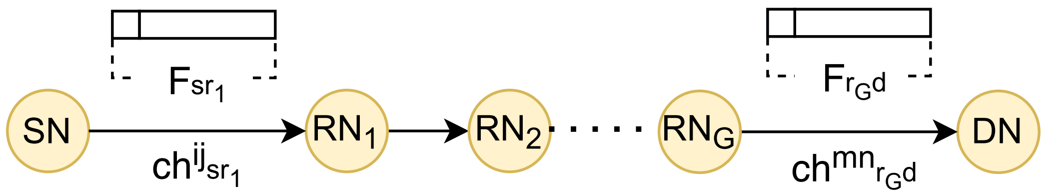



Figure 1.1: A relay network with SN, multiple RNs, and DN. The RNs perform multi-band communication. The forwarded data frame ($F$) is also exhibited in the illustration.

notations.

On the other hand, in multi-band relay network communication, as $\mathrm{RN_k}$ receives packets from SN or a preceding $\mathrm{RN_{k-1}}$, it may start dispatching packets to DN or a subsequent $\mathrm{RN_{k+1}}$ using another available $m^{\text{th}}$ channel of $n^{\text{th}}$ band, $\mathrm{ch}_{rd}^{mn}$, even before the data transmission from SN (or $\mathrm{RN_{k-1}}$) to $\mathrm{RN_k}$ is completed. Multiband relay network can ensure efficient spectrum usage (with the usage of Cognitive Radio (CR) discussed in Section 1.2.2) by simultaneously increasing coverage and reducing data transmission delay.

This concept could be further extended to splitting a single data frame into multiple sub-packets at the RN and distributing them over multiple sensed available channels, ensuring further efficiency in spectrum usage while increasing coverage and reducing data transmission delay. In this work, the RN re-transmits the entire frame instead of breaking it into sub-packets.

### 1.1.4 Network Approaches

As solutions to the latency issue in relay-based communications, both centralized and distributed approaches are introduced, in order to optimize channel allocation over multiple

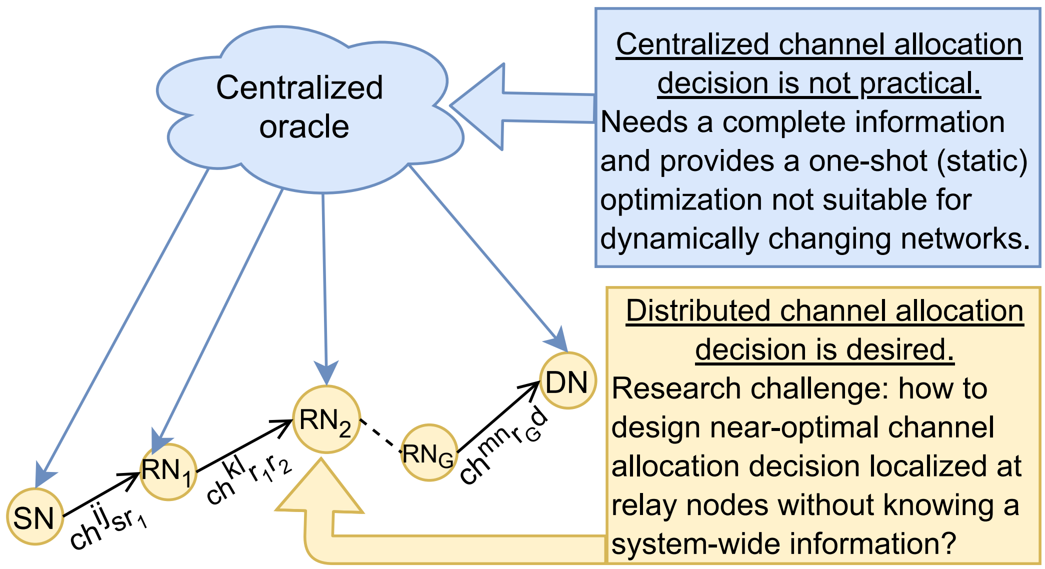



Figure 1.2: Centralized vs distributed resource selection for relays.

heterogeneous bands.

1. *Centrally Controlled Network.* As depicted in Fig. 1.2, when a centralized oracle makes decisions for all network entities, including the RNs, it requires information on the complete network to provide a one-time (non-adaptive) solution. Such a decision-making process requires a non-deterministic amount of time with a significant increase in the number of nodes. Moreover, such a solution is clearly unsuitable for rapidly-evolving channel environments. This results in a more complex, time-consuming, and static optimization, which is not practically viable for relay networks. Therefore, designing a practical band/channel selection method that minimizes the adverse impact on QoS parameters (e.g., communication delay, throughput, available power/energy, available buffer) in relay-based communication emerges as a key research problem in this thesis.

2. *Distributed Control Network.* The desired approach, therefore, hinges upon designing a distributed, online decision-making technique that permits RNs to localize heterogeneous bands/channel selection-based on the prevalent channel and traffic conditions. Such a localized distributed decision-making process is naturally lightweight (i.e., less complex), prompt and reliable, thereby improving the overall delay (from SN to DN) and throughput.

### 1.1.5 Graph Theory

In mathematics, graph theory [3] is the study of graphs, which are mathematical structures used to model pairwise relations between objects in a wide environment. Graphs-based models provide a high-level representation to complex mathematical structures of a complicated scenario. A graph, in this context, consists of vertices (also called nodes or points) which are connected by edges (also called links or lines). Edges are ordered pairs of vertices (that is, an edge is associated with two distinct vertices). Types of graphs: Directed, Un-Directed, Colored, Bipartite, Circular etc.

With regards to the centralized oracle approach, in this thesis, a graph theory-based solution is proposed. To be precise, the problem is modeled as a bipartite graph and solved using tools from optimization theory [4]. Graph theory is ultimately a study of relationships which provides a helpful tool to clarify and evaluate complete dynamics of the problem, hence providing a brief interpretation of a convoluted network.

### 1.1.6 Game Theory

Game theory [5] studies strategic interaction between individuals/nodes in situations, called games. Game theory provides a framework for players in interacting environments. Types of games: Cooperative/Non-Cooperative, Constant/Zero/Non-zero sum, Simultaneous/Sequential move games.

For a distributed approach with rapid decisions and performance guarantee, a sequential game-based technique is also proposed so as to solve the problem that satisfies all the related constraints in a well defined structure. Game theory provides a proper framework for understanding choices in the situations having competing players hence making this approach more appealing to explore and study in this scenario. In proposed sequential game with perfect information (not to be confused with complete information, which could be observed in the case of the centralized oracle), the RNs act as players.

### 1.1.7 Nash Equilibrium

In game theory, in order to find the best strategy for each player, the theorem of Nash Equilibrium is used [6]. Nash Equilibrium is an ideology where each device (player) has some or all the information of the corresponding network and there are no other parameter adjustments (moves/strategy) left that can enhance its performance (utility). To choose its next strategy, each player first looks for the network (other devices) information available to it and then opts for the strategy that can enhance its performance (a step forward).

In other words, it can be said that each player chooses a strategy from which it will never deviate (as any other strategy will not enhance its performance) so long as it cannot go a step higher. Nash Equilibrium is also said to be the most "dominant strategy" leading towards better results out of all other possible moves. Nash strategy is not always the most optimal strategy but is actually the best one. In this thesis, the model-specific Nash Equilibrium analysis is also provided for proposed game theory model.

### 1.1.8 Markov Decision Process (MDP) and Reinforcement Learning

A Markov process defines all the states of a certain agent, which is actively involved in coordinating with environment rewards against certain actions. The agent being in a certain state, $S$, takes an action, $A$. According to the action taken by the agent, the environment now assigns a particular reward, $R$, against that action and also decides the next state, $S'$ for the agent to move onto. With this continuous interactions, the RN acting as the agent keeps learning the best states and actions for itself. The actions are taken on the basis of a certain policy that could be either totally random, or greedy or epsilon-greedy ($\epsilon$-greedy).

The transition from one state to the next has certain probabilities. These probabilities are known as transition probabilities ($p$) and can be written in the form of transition tables [7]. The latency problem is reformulated as a finite MDP reinforcement learning model, which consists of a stochastic 4-tuple of states, actions, rewards and transition probabilities $(S, A, R, p)$ so as to handle the channel dynamic and user mobility.

## 1.2 Principal Assumptions

In this section, a brief overview of principal assumptions is provided, which are presumed throughout the thesis. These assumptions are primary design ideas which are worth considering for any future practical implementation of the said problem, algorithms and approaches.

### 1.2.1 SNR Estimation

Before re-transmission of any data packet, it is assumed that all the nodes (or just the CN in case of oracle approach) perform SNR estimation in order to check the quality of the channel/frequency. There are multiple methods of performing SNR estimation such as [8, 9, 10]. The main idea is to send a pilot signal to the destined device and to get a reply on bit error rate or packet error rate at the receiver and then to estimate the SNR of the particular channel using mean square error approach.

### 1.2.2 CR and Spectrum Sensing

Wireless radio spectrum is a finite resource, making it a challenge to accommodate the continuously increasing number of users and the demand for high data rate. This 5G+ generation of wireless communication consequently urges an efficient usage of the spectrum. An intelligent radio with adaptive features and advanced network technology seems to be the solution. Here, available free channels can be automatically detected and utilized, instead of occupying specific portion of spectrum and keeping it under-utilized. CR serves as an example of such an advanced communication module, where the well-known technique of spectrum sensing is adopted by the secondary users to efficiently use the under-utilized licensed spectrum of the primary users [11]. In this thesis, the idea of spectrum sensing and utilizing the free spectrum for communication is investigated, allowing for an efficient use of the spectrum.

### 1.2.3 Allocation of RN/s

This assumption is worth mentioning and to be considered, as this thesis work can be confused between band/channel allocation approaches and RN allocation algorithms. It is very explicitly assumed throughout the thesis that before SN starts any packet transmission, SN itself, fellow RN/s and DN; all of them are already aware of each other for current packet/s transmission.

### 1.2.4 Buffer Limitation

Another important subject of this thesis is the buffer limitation which, in addition of being an assumption, is also an important conceptual background. In intended multi-band communication system, where a node is capable of transceiving data packet on any two distinct frequencies, there is a high probability of the fact that the data rates are not going to be the same; having miss-matched SNR estimations. In brief, the channel having good SNR estimation is capable of transceiving on higher data rate where as, the channel with poor SNR conditions can transceive on lesser data rates only. Such kind of communication needs buffer to store data before or after packet gets re-transmitted. This buffer is assumed to be of a limited size, throughout the thesis.

## 1.3 High Level Problem Description

Although emerging relay-based communication networks boast of increased transmission coverage and capacity, they currently lack schemes that optimize the simultaneous utilization of the best-quality channels, especially when such channels belong to heterogeneous frequency bands. With advanced multi-band capable hardware devices, there is now need of detailed theoretical, mathematical and simulations-based analysis of reduced packet latency in multi-band relay networks.

In this thesis, the pros and cons of relay transmission and multi-band communication are cumulatively studied, hence proposing both centralized and distributed approaches to achieve reduced packet relay latency. This time-sensitive problem needs to be solved quickly (fast decisions) rather than optimally. In centralized optimization approach, as the number of RNs increase, the optimal solution is difficult to determine. On the other hand, if the distributed approach is used, the data communication latency is significantly reduced using local feasible solutions. Hence the main objective of this thesis is to utilize both relay network and multi-band systems in order to reduce the total packet latency sent from SN via RN/s till DN.

## 1.4 Thesis Outline

The rest of the thesis comprises of brief literature background in Chapter 2, where the summary of related work is provided along with their contributions and drawbacks.

In Chapter 3, a relay-assisted network is introduced, leveraging multiple frequency bands simultaneously. First, an optimal heterogeneous band/channel allocation strategy is formulated. Then, the in-feasibility of this centralized approach is elucidated, resulting in increased latency. As a solution, a heuristic-based, distributed channel allocation approach is proposed whereby each RN is able to make localized decision regarding the best channel allocation over heterogeneous bands. Adaptability analysis and computer-based simulation results demonstrate that proposal outperforms the conventional method with much reduced relay latency.

In Chapter 4 a centralized oracle is presented, based on a bipartite graph model for the multi-band multi-channel allocation to RNs, and indicate why this is a non-viable solution for deployment. Therefore, with reference to previously proposed greedy approach, a better online, distributed multi-band/channel allocation strategy is explored by proposing a sequential game-theoretic algorithm. Simulation results demonstrate that proposed game-theoretic approach significantly outperforms the traditional distributed and centralized methods.

In chapter 5, the proposed game-theoretic model is further analyzed in depth using model-specific Nash Equilibrium. In game theory the players need to find equilibrium among each other so as to maintain the feasible solution with mutual arrangements. In proposed sequential game model, such equilibrium is found by letting the players play sub-games in backward iteration. Such method is known to be sub-game perfect equilibrium using backward induction. In chapter 5, the detailed analysis of said approach is provided and explained using illustrative examples.

Chapter 6 has details on machine learning-based model, proposed to handle the dynamic channel environment with mobility of user equipment. Previously proposed algorithms assume much stable channel conditions. Where as, the reinforcement learning-based algorithm is designed to learn from probable states and action in order to eventually learn the optimum distributed solutions.

## 1.5 List of Publications

**Published**

1. B. Mughal, Z. M. Fadlullah and S. Ikki, "Centralized Versus Heuristic-Based Distributed Channel Allocation to Minimize Packet Transmission Delay for Multiband Relay Networks," *IEEE Networking Letters*, vol. 2, no. 4, pp. 180-184, Dec. 2020.

2. B. Mughal, Z. Fadlullah, M. M. Fouda and S. Ikki, "Allocation Schemes for Relay Communications: A Multi-Band Multi-Channel Approach Using Game Theory," *IEEE Sensors Letters*, Accepted, 2022.

**Submitted**

3. B. Mughal, Z. M. Fadlullah, M. M. Fouda and S. Ikki, "Applying Game Theory to Relay Resource Selection in Hybrid-band Wireless Systems," Submitted to *IEEE Sensors Journal*, Nov. 2021.

4. B. Mughal, Z. M. Fadlullah, M. M. Fouda and S. Ikki, "Optimizing Packet Forwarding Performance in Multi-Band Relay Networks via Customized Reinforcement Learning," Submitted to *IEEE Transactions on Vehicular Technology*, Dec 2021.

# Chapter 2

# Literature Review

## 2.1 Introduction

Reducing communication delay is an active area of research in relay-based networks. Relay communication using a multi-band multi-channel approach has been studied in [12, 13, 14, 15, 16, 17, 18] with the aim of reducing end-to-end packet latency.

## 2.2 Prior Work

A multi-hop network was proposed in [12], where a centralized optimization problem was employed to offer scheduling decisions in terms of routing, channel allocation and link scheduling. This work considers the fact that uncertain channel availability results in uncertain packet scheduling, which not only increases latency but also heavily overloads the system. Such centrally-controlled relay multi-band systems can only be used for smaller networks, where it can be assumed that channel conditions for a specific range would be the same (as sensed by the Central Node (CN)) and would remain constant for a certain period of time.

A multi-band transmission-based relay model is proposed in [13]. The proposed model is for Wireless Local Area Network (WLAN) systems and aims to reduce packet latency. This work carries out relay transmission and frame reception at the same time by reading the truncated header instead of complete header. To re-phrase it, while receiving the data packet, the RN relays (re-transmits) data packet to the next node right after reading the truncated header instead of reading it completely.

Such strategy is called as truncated decode and forward scheme. In order to obtain the information for judging whether or not relay transmission is required, the RN receives the header of a frame on one frequency band. If relay transmission is required, the RN starts the relay transmission on another frequency channel while the RN continues to receive the frame [19]. However, reading partial header, increases the chances of frame error as, it can degrade the de-coding performance. This might result in packet re-transmission, hence increasing the packet latency.

In [14], the authors provided practical experimental results for a multi-band WLAN system. This system finds scattered and unused spectral resources across multiple bands and uses them for frame transmission. Therefore, utilizing multi-band WLAN can improve the total spectral efficiency when compared to the legacy single-band WLAN.

Furthermore, to effectively utilize spectral resources, the algorithm judges, based on the prediction of an idle/busy channel, whether the transmitter should immediately transmit using a single band or wait for a while to make multi-band simultaneous transmission. This work does not utilize the relay technology and can therefore only be used for lesser range networks.

In [15], the authors proposed a lightweight algorithm that considers appropriate channel usage for receiving/sending data, based on Signal to Interference Noise Ratio (SINR), with the ultimate aim of avoiding data loss caused by the difference between SN and RN rates. This work adopts spatial re-usage by adjusting a threshold of acceptable interference.

In this algorithm, the channels are selected such that the receiving and sending rates are maximized while the transmission delay is minimized. In the algorithm, the SINR of each channel is first calculated for SN to RN and RN to DN, and the modulation and coding scheme is chosen by looking up an appropriate scheme corresponding to the SINR. This work lacks a strategy that handles the case where better SINR channels are unavailable.

### 2.2.1 Algorithms based on Awareness

In [16], a similar approach is proposed but there is no handling of any specific cases where low SNR channel and high SNR channel are available. Researchers in [17] presented a multi-channel allocation algorithm, based on Congestion-Aware (CA) in an attempt to reduce the interference between nodes and extend their longevity. The algorithm determines the degree of congestion degree across different channels by comparing the transmission capacity of each of their nodes, ultimately selecting the most suitable channel for data transmission. This algorithm reduces the number of channel selection conflicts and improves the throughput of the whole network as much as possible.

A channel ranking algorithm was proposed in [18], in which all nodes prioritize available channels on the basis of their available channel properties. Afterwards, a distributed channel allocation algorithm is implemented such that each node can choose a suitable channel, based on both: its residual energy and channel ranks.

Finally, the spectrum-sensing and sleep duration are optimized together in order to satisfy energy consumption constraints and increase the normalized throughput, simultaneously. This approach requires no global information or any central coordination, unlike other former approaches. Here, each node can work in an absolute distributed manner, based only on its own local knowledge.

Furthermore, theoretical analysis and extensive simulations have validated that, upon applying this solution to the IoT network: (*i*) each node can be allocated to a proper channel, based on the residual energy to balance the lifetime; (*ii*) the network can rapidly converge to a collision-free transmission through each node's learning ability during the process of the distributed channel allocation; and (*iii*) the network throughput is further improved via dynamic time slot optimization. Here, this work is referred as a Power-Aware Congestion-Aware (PACA) algorithm. Later the proposed models are compared with these conventional CA and PACA approaches.

### 2.2.2 Graph and Game Theory Literature:

Graph theory, being a widely used representation approach for complex expanded networks, and game theory, being the most preferred technique for handling environments with competing nodes/players, are studied in literature [3, 20, 21, 22, 23] for solving typical resource allocation problems in wireless communication networks. In spite of the literature being enriched with game and graph theory work for allocation of various network resources, it aims for increased sum rate only and lacks at using game and graph theory with aim of solving the problems for increased data packet latency in relay network and buffer overflow situations of multi-band network.

Proposed work on graph and game theory, not only aims to reduce data packet latency but also shows its effect on the effective throughput of the entire network. Hence it can be said that now is a much more opportune time to solve the added packet latency and buffer overflow problem for multi-band relay network using advanced techniques like game and graph theory.

### 2.2.3 ML based Algorithms

While such distributed methods typically rely on pre-established rules, the highly dynamic network conditions and mobility makes it difficult to manually design all the rules and incorporate them to the decision algorithm in a hard-coded manner. To address this issue, machine learning techniques have been extensively considered in the literature to tackle similar computational problems in various communication network scenarios. For instance, reinforcement learning, have been employed in different types of Cognitive Radio

Network (CRN) for adaptive access of the dynamic spectrum.

Felice *et al.* presented a survey as well as a proof-of-concept of reinforcement learning-assisted spectrum management for CRNs [24]. The categories of spectrum sensing problem was further investigated in [25] that revealed that the sequential multi-channel selection in CRNs leads to sensing-order and stopping rule problems.

On the other hand, periodic and channel-specific spectrum sensing was identified to be a Cooperative Sensing Scheduling (CSS) problem leading to under-utilization of the available spectrum. However, this work pointed out that a distributed channel allocation in a CRN setting is recommended; however, is challenging due to the aforementioned problems. However, none of the aforementioned work investigated the coexistence of multiple frequency bands for the channel allocation problem let alone taking into account the relay-based topology and user/relay movement in the emerging networks.

The work in [26] demonstrated how the different frequency bands having various propagation properties, particularly in the mmWave bands, affects the hand-off mechanism by the need to select the best base-station via the best possible frequency band. The UE-base-station hand-off policy is developed as a MDP to achieve the optimal hand-off decision by considering the remaining bandwidth of the serving base station, link conditions, and the user's maximum pay per connection budget.

### 2.2.4 Lessons Learned

The lessons learned from the literature review can be summarized as follows to derive the motivation behind work done in this paper. The centralized solutions are typically computationally hard and non-deterministic polynomial time in nature, and the updated channel conditions are usually not considered to avoid a highly complex scenario.

On the other hand, the management of limited buffer at RNs is also not considered in many existing research work that can have a detrimental effect on the packet latency in the relay network. Furthermore, corner cases involving high and low SNR impacting the throughput and delay performance of RNs are also not taken into account in the existing researches. Hence, a system model to address these various aspects is required that is presented in this thesis.

## 2.3 Literature Shortcomings

1. Once an advanced system with multi-band capability is considered, there is a difference in data rates of received and re-transmitted packet. In literature, for such a scenario, the detailed mathematical analysis for lower and higher data rates is not given.

2. The literature is more focused on finding the optimal solution to this time-sensitive problem of reducing latency instead of fast and feasible solutions.

3. Once the received and forwarded data rates are different, buffer limitations are not handled efficiently.

4. Many authors in literature have focused on increasing throughput of the multi-band system, ignoring the time delay getting added because of added multi-hop nodes.

## 2.4 Thesis Contributions and Learning

1. In this thesis mathematical equations are provided for both: the cases where higher or lower SNR channels are available for packet re-transmission.

2. This work focuses on finding feasible and faster solution to the problem of reduced latency instead of finding much delayed optimal solutions.

3. This thesis also designs the oracle solutions to the problem which are then proved inefficient in terms of reducing latency.

4. While focusing on reducing time delay for advanced multi-band relay network, this work also aims to improve effective throughput of the system.

5. This work is based on designing, formulating and solving the latency problem as an optimization theory issue.

6. Along with keeping QoS and power usage as a basis of design, this work also uses advance technologies like game theory, bipartite graph model and the concepts of Nash Equilibrium in order to reduce latency further.

## 2.5 Conclusion

From literature review it can be seen that there is a need of analysing the multi-band relay systems in detail, with focus on managing the multi-rate scenarios where, the higher SNR channels are available or only lesser SNR channels are available. Moreover, there is a need to utilize less complicated and advance techniques to allocate the available channels efficiently. There is also a space to fill in proving that multi-band relay systems are to be handled distributively and efficiently, instead of optimally and centrally.

# Chapter 3

# Centralized Versus Heuristic-based Distributed Channel Allocation to Minimize Packet Transmission Delay for Multiband Relay Networks

**Motivation:** In case where, multi-band relay network is not well designed for efficient channel allocations, it might result having increased packet latency and also increased spectrum usage. Such a system requires a less complicated but efficient approach to assign the best possible channels in order to train the network for reduced latency, instead of optimal channel allocation.

## 3.1 Introduction:

In this chapter, an optimal channel allocation problem is formulated, over heterogeneous frequency bands to multiple relays. Then, it is demonstrated that solving this optimization problem requires centralized coordination and high computation which may not be possible in RNs that may be resource-constrained (e.g., Device 2 Device (D2D) RNs) and subject to small-scale channel fading.

As a solution, a heuristics-based distributed approach is envisioned, where each RN can make its own decision locally to select the best possible channel from a pool of frequency bands. Computer-based simulation results demonstrate that proposal outperforms the centralized baseline in terms of execution time and the conventional methods with much reduced relay latency.

The remainder of the chapter is structured as follows. Section 3.2 provides centralized design of the problem for optimal solution. This section provides formal problem formulation of the optimal channel allocation to RNs, leveraging multiple bands in which centralized coordination is required. Then brief introduction to conventional scheme is provided in Section 3.3. A distributed solution is presented in Section 3.4. The performance of proposal is evaluated in Section 3.5. Finally, concluding remarks are provided in Section 6.6. The details and explanation of some basic notations used throughout the thesis can be found in Table 6.1.

Table 3.1: Notations Table

| Notations | Description |
|---|---|
| $F_{sr}$ | Data frame sent from SN to RN |
| $F_{rd}$ | Data frame from RN to DN |
| $t_o$ | Time at which packet transmission starts from SN |
| $t_{ir}$ | Time at which RN starts receiving $F_{sr}$ |
| $t_{or}$ | Time at which RN schedules the data frame for re-transmitting it to DN ($F_{rd}$) |
| $T_h$ | Time span of the header of the packet. |
| $T_{sr}$ | Time span of $F_{sr}$ |
| $T_{rd}$ | Time span of $F_{rd}$ |
| $T_{tot}$ | Total time span from the instant at which RN starts receiving data frame from SN ($F_{sr}$) till the time instant, DN completely receives the re-transmitted data frame ($F_{rd}$) from RN. |
| $D_{sr}$ | Rate at which SN transmits data frame to RN ($F_{sr}$) |
| $D_{rd}$ | Rate at which RN re-transmits data frame to DN ($F_{rd}$) |
| $ch_{sr}^{ij}$ | i-th frequency channel of j-th frequency band used by SN to transmit $F_{sr}$ to RN |
| $ch_{rd}^{mn}$ | m-th frequency channel of n-th frequency band used by RN to re-transmit $F_{rd}$ to DN |
| $\text{SNR}_{sr}$ | Estimated SNR of the channel used by SN to transmit $F_{sr}$ to RN |
| $\text{SNR}_{rd}$ | Estimated SNR of the channel used by RN to re-transmit $F_{rd}$ to DN |

## 3.2 Centralized Approach

In this section, the problem of channel allocation to multiband RNs with the coordination of a centralized entity. Let G denote the number of possible RNs (RNs). Let Q represent the number of sequences of free channels over heterogeneous frequency bands, that could possibly be assigned to these RNs. Let **SR**, **BO**, **SNR**, $\mathbf{t_{or}}$, $\mathbf{T_{rd}}$ be matrices, each with $Q \times G$ dimension. Here, **SR** consists of all possible channel sequences and **BO** is the corresponding buffer overflow matrix. $t_{or}$ represents the time at which packet is being scheduled while $T_{rd}$ denotes the corresponding time span of the packets. The optimization problem is modeled as a resource allocation problem as follows:

$$\min \quad \sum_{q=1}^{Q} X_q(t_{or,qG} + T_{rd,qG}), \qquad \text{s.t.} \tag{3.1a}$$

$$\sum_{g=1}^{G} X_q BO_{qg} = 0, \quad \forall q, \tag{3.1b}$$

$$X_q SR_{qg} \neq X_q SR_{q(g+1)} \quad \forall q, g, \tag{3.1c}$$

$$\sum_{q=1}^{Q} X_q = 1 \quad \forall q, g, \tag{3.1d}$$

$$\sum_{g=1}^{G} X_q \mathrm{SNR}_{qg} \geq \mathrm{SNR}_{min} \quad \forall q, \tag{3.1e}$$

$$X_q \in \{0, 1\}, \tag{3.1f}$$

where, $X_q$ is a column vector of unknown binary decision variables. The details of objective function and constraints are summarized below:

1. The objective function (3.1a) of this optimization problem is to minimize the total end-to-end packet transmission time $(t_e - t_o)$, where $t_o$ is the time at which packet transmission starts from SN and $t_e$ is the time at which the packet transmission finishes at DN, calculated as the sum of $t_{or,qG}$ and $T_{rd,qG}$, which denote the time at which $G^{\text{th}}$ RN transmits the packet out to DN and the time span of this packet, respectively.

2. Constraint (3.1b) enforces the selection of a sequence having buffer overflow=0 at each RN (i.e., $BO_{r_1} = BO_{r_2} = \cdots = BO_{r_G} = 0$).

3. Constraint (3.1c) selects different channels of differing or identical bands at each RN, allowing for the use of multiband communication, i.e., $ch_{sr_1}^{ij} \neq ch_{r_1r_2}^{kl} \neq ... \neq ch_{r_Gd}^{mn}$.

4. Constraint (3.1d) enforces the selection of at least a single channel sequence.

5. (3.1e) is a QoS constraint keeping SNR higher than the minimum acceptable $\text{SNR}_{\text{min}}$. This is aimed to reduce the probability of bit error rate [27].

The above formulation is an integer linear programming model, optimal value of which can be found using the branch and bound algorithm [28]. However, this problem is NP-hard (non-deterministic polynomial time hard). The drawbacks of solving this problem centrally in the relay-based system model are as follows:

- Reducing latency in the considered relay network is a time-sensitive issue. While the integer programming problem could be optimally solved, as the number of RNs increases, the optimizer takes non-deterministic time to reach the solution. This adds more latency to the data communication. Moreover, during the time to derive the optimal solution, the channel conditions are likely to change for which the computed optimal solution is no longer valid.

- Such conventional, centrally-controlled networks depend on channel availability information, i.e., at time $t_1$, during the start of running algorithm. Once the optimal solution is derived at time $t_n$, the channel conditions are likely to change that renders the estimated optimal solution invalid.

Therefore, it is critical for the individual RNs to make localized decisions, without the network-wide perfect information on channels availability for all RNs, for assigning the best channels from multiple bands. A distributed solution is explored in the following section.

## 3.3 Conventional Decode and Forward (DF)

Conventionally, in DF scheme, the RN has to wait for the receiving channel to get free so that it can forward the relayed data packet to the next node. This increases packet transmission latency as shown in packet timing diagram of DF in Fig. 3.1.

In next section a multi-band relay communication model is proposed using a distributed greedy algorithm to reduce the added latency of relay communication network. The concept is then extended in order to handle the two possible cases that can be faced by a RN in

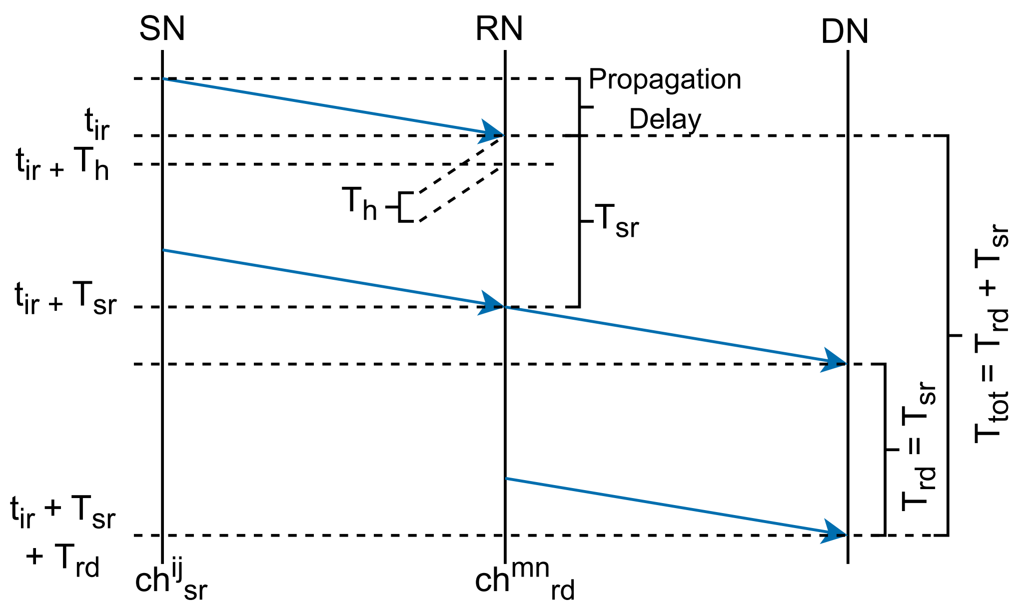



Figure 3.1: Packet timing diagram of DF system.

order to forward the data packet to the next node/destination: availability of higher SNR channel or a lower SNR channel. Once a SN transmits a packet to a RN at specific data rate, the RN, using multi-band communication, can forward the data packet to the next node/destination, at a different channel using an alternate band (hence using a different data rate). This difference in data rates requires an available buffer to store the data.

As it can be noticed from the flowchart steps illustrated in Fig. 3.2, in customized greedy heuristic, upon receiving the data packet from the SN at $D_{sr}$, the RN checks if it has lesser or greater SNR channel available in the same or different band. Then, according to its local channel availability, the RN chooses the forwarded packet's data rate and scheduled time.

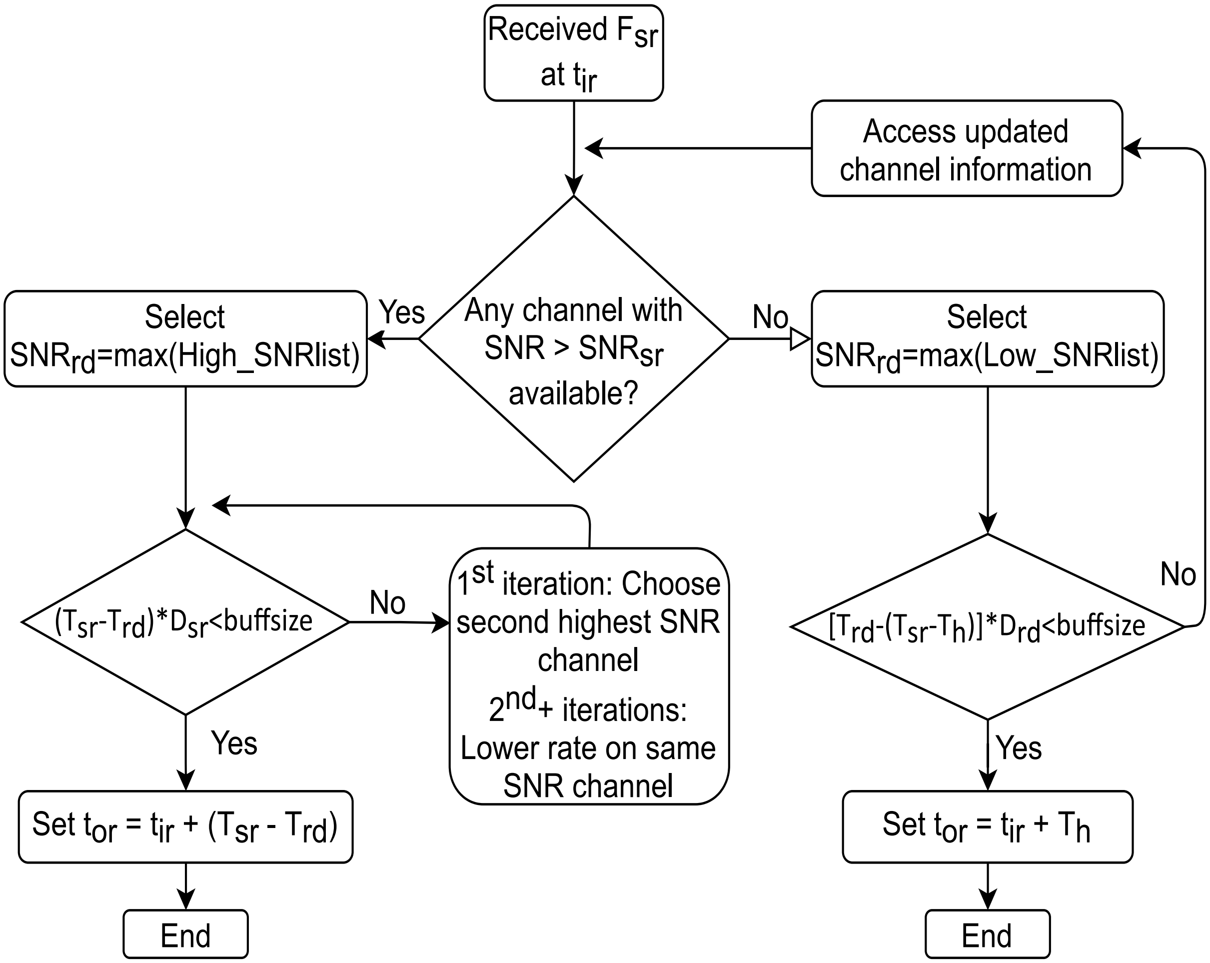



Figure 3.2: Steps of the proposed greedy heuristics.

## 3.4 Considered System Model and Proposed Distributed Channel Allocation Method Leveraging Multiple Bands

### 3.4.1 System Model

For simplicity, in this section, a system model with a single RN is considered, based on IEEE 802.11ax specifications [29]. The objective is to make quick, localized decisions using a heuristic-based distributed approach. This work also aims for an efficient spectrum utilization, reducing transmission latency with multiband communication and increasing coverage with the RNs.

Proposed heuristic algorithm is designed in a lightweight manner to run at each resource-constrained RN. It is assumed that before starting re-transmission, each RN accesses its free channels database and performs SNR channel estimation [30] on all available channels across multiple bands. At $t_{ir}$, the RN starts receiving the data frame $F_{sr}$ from the SN, and starts buffering it.

Next, the RN needs to re-transmit the data frame ($F_{rd}$) to the DN. Then, the RN computes time instant, $t_{or}$, to schedule $F_{rd}$ out to the DN. While finding the best suitable value of $t_{or}$, the algorithm adjusts itself according to the size of the available buffer and the estimated SNR of the available channels. Next, the algorithm tends to keep any channel occupied for the shortest possible duration of time, and achieve the minimum possible value for the total SN to DN transmission time, $T_{tot}$.

Table 3.2: Considered modulation and coding scheme (MCS). Notations used: Mod = Modulation, CR = Coding Rate, Cb/sub = Coded bits per subcarrier, Cb/sym = Coded bits per OFDM (orthogonal frequency-division multiplexing) symbol, Db/sym = Data bits per OFDM symbol, DR = Data Rate, BW = bandwidth.

| **MCS** | **Mod** | **CR** | **Cb/ sub** | **Cb/ sym** | **Db/ sym** | **DR (Mbps) Channel BW (MHz)** | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | **20** | **10** | **5** |
| 0 | BPSK | 1/2 | 1 | 48 | 24 | 6 | 3 | 1.5 |
| 1 | BPSK | 3/4 | 1 | 48 | 36 | 9 | 4.5 | 2.25 |
| 2 | QPSK | 1/2 | 2 | 96 | 48 | 12 | 6 | 3 |
| 3 | QPSK | 3/4 | 2 | 96 | 72 | 18 | 9 | 4.5 |
| 4 | 16QAM | 1/2 | 4 | 192 | 96 | 24 | 12 | 6 |
| 5 | 16QAM | 3/4 | 4 | 192 | 144 | 36 | 18 | 9 |
| 6 | 64QAM | 2/3 | 6 | 288 | 192 | 48 | 24 | 12 |
| 7 | 64QAM | 3/4 | 6 | 288 | 216 | 54 | 27 | 13.5 |

Table 3.3: MCS-SNR Ranges

| MCS | SNR Ranges |
|---|---|
| 0 | SNR $\leq 0$ |
| 1 | $0 \leq$ SNR $< 3$ |
| 2 | $3 \leq$ SNR $< 6$ |
| 3 | $6 \leq$ SNR $< 9$ |
| 4 | $9 \leq$ SNR $< 12$ |
| 5 | $12 \leq$ SNR $< 15$ |
| 6 | $15 \leq$ SNR $< 18$ |
| 7 | SNR $> 18$ |

The SN transmits $F_{sr}$ to the RN at a specific data rate, $\mathcal{D}_{sr}$, by using the modulation and coding scheme (MCS) [31] in Table 3.2. The same case is applicable with the re-transmission of $F_{rd}$ at rate $\mathcal{D}_{rd}$. Table 3.2 is used in accordance with the SNR condition of the available transmission channel. If the RN finds any free channel with higher SNR value than SN-RN channel's SNR, RN is able to transmit at a higher rate. If there are channels with lower SNR only, the RN transmits at a lower rate. These cases are discussed in the remainder of the section.

Proposed algorithm assigns rates according to the SNR estimation ranges [15], Table 3.3. Because the rate and time are inversely proportional to each other, it is vital to note that the efficient selection of $\mathcal{D}_{sr}$ and $\mathcal{D}_{rd}$ implies increase or reduction in total time taken ($T_{tot}$) for the complete data frame transmission from SN to DN.

As there is a difference in $T_{rd}$, $T_{sr}$, and in $t_{ir}, t_{or}$, a buffer is needed at the RN to keep data before/after and during data re-transmission. In addition, it is assumed that the RN performs spectrum sensing in regular intervals and maintains the database of sensed available channels locally.

There are two main assumptions throughout this chapter. First, the RN always receives and decodes the entire header of the received packet from SN. This implies that the RN will start re-transmitting the same packet only after $(t_{ir} + T_h)$, which is the time taken to send/receive the header of the packet. Second, all the processing and propagation delays are assumed to be negligible and are not included in the total time calculation, $T_{tot}$.

### 3.4.2 Proposed Heuristic Algorithm

Fig. 3.2 and Algorithm 1 shows the steps of proposed heuristic algorithm. Upon receiving $F_{sr}$, the RN scans its database containing the list of SNR estimation results for the sensed free channels, $\text{SNR}_{list}$. Thus, proposed algorithm encounters two scenarios:

In the **first** (also the most favourable) case, the RN finds out that one or more free channels have $\text{SNR} > \text{SNR}_{sr}$. This means that $F_{rd}$ could be transmitted to DN at a higher rate, $\mathcal{D}_{rd}$. In other words, $\mathcal{D}_{rd} > \mathcal{D}_{sr}$ and $T_{rd} < T_{sr}$ as shown in Fig. 3.3.

The **second** case occurs when only lower SNR channels are available, i.e., $\text{SNR} < \text{SNR}_{sr}$ for packet re-forwarding. If this case occurs, the RN might result in sending $F_{rd}$ with a lower rate ($\mathcal{D}_{rd} < \mathcal{D}_{sr}$), implying $T_{rd} < T_{sr}$, as shown in Fig. 3.4.

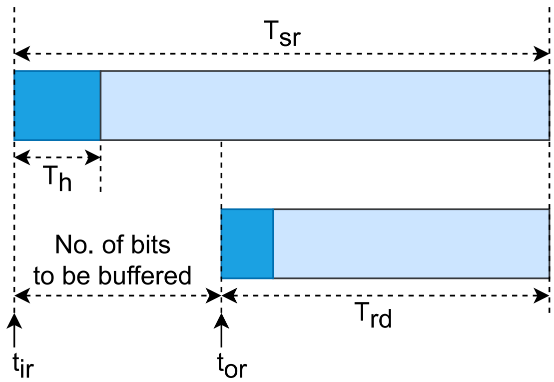



Figure 3.3: Frame diagram for the case of high SNR channel available.

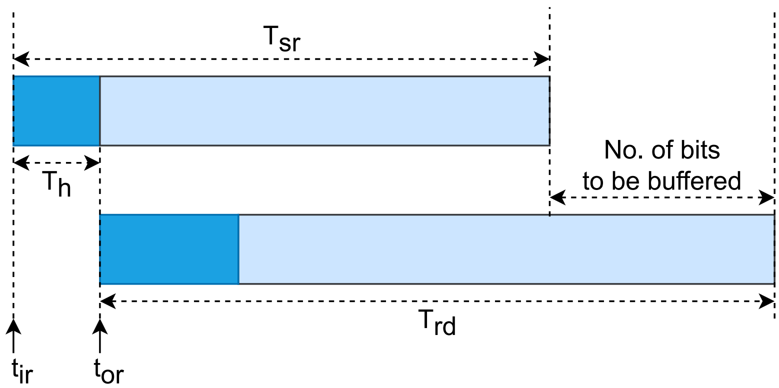



Figure 3.4: Frame diagram for the case of low SNR channel available.

Moreover, before selecting $t_{or}$ in both cases, the respective Buffer Overflow Condition (BOC) is verified and the algorithm accordingly adjusts itself. BOC refers to the condition in which the RN checks if the buffer-size can hold the data bits before, during and/or after the re-transmission of $F_{rd}$. The occurrence of these two cases and proposed algorithm's adaptability for both cases is explained in the remainder of the section.

**Algorithm 1** Pseudocode for Proposed Algorithm at each RN.

**Input:** $F_{sr}, \mathcal{D}_{sr}$

**Output:** $t_{or}$

```
flag=0, true=0, bufferoverflow=1
while true==0 do
    if better SNR channels exist, flag≥1
    if only low SNR channels available, flag=0
    if flag≥1 then
        while bufferoverflow≥1 do
            Calculate buffer needed for chosen channel
            if bufferneeded<buffersize then
                bufferoverflow=0, true=1
                output t_or
            end
            if  then
                Select 2nd higher SNR channel, Update D_rd
            else
                Lower the data rate on same chosen highest SNR channel, Update D_rd
            end
        end
    else
        while bufferoverflow==1 do
            Calculate buffer needed for chosen channel
            if bufferneeded<buffersize then
                bufferoverflow=0, true=1
                output t_or
            end
            if then
                Get updated channel information
                bufferoverflow=0
            else
                true=1, bufferoverflow=0, Send using DF
            end
        end
    end
end
```

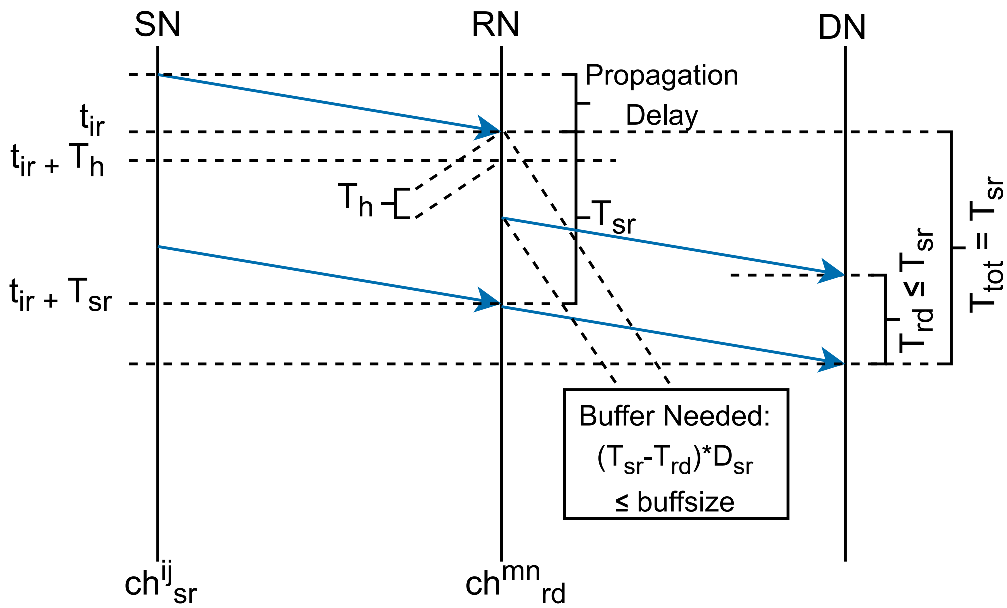



Figure 3.5: Packet timing diagram for the case where higher SNR channels are available.

### 3.4.3 Adaptability Analysis of Proposed Heuristic Algorithm

In this section, both cases of lower SNR channel available and higher SNR channel available are discussing in detail along with their detailed timing diagrams and corresponding equations.

#### CASE 1 - Higher SNR channels available

As discussed earlier, the case of a higher SNR channel available is the most preferred case and its detailed timing diagram is depicted in Fig. 3.5. In this case, the RN has better SNR channel(s) available than $\text{SNR}_{sr}$ for $F_{rd}$ re-transmission. This case is the most preferable scenario, leading to the shortest value of $T_{tot}$ as shown in Fig. 3.3. As the RN finds any channel with higher SNR than $\text{SNR}_{sr}$, it selects that channel and checks the BOC:

$$(T_{sr} - T_{rd})\mathcal{D}_{sr} < \text{buffer-size} \tag{3.2}$$

Here, $T_{rd}$ denotes the time span of $F_{rd}$ if it is to be sent on this chosen SNR with its corresponding rate. BOC checks if the RN has buffer-size available for storing $(T_{sr} - T_{rd})\mathcal{D}_{sr}$ number of bits of $F_{sr}$ packets before the re-transmission of $F_{rd}$ starts (during the reception of $F_{sr}$). In this case, $F_{rd}$ can be sent at a higher rate than $\mathcal{D}_{rd} > \mathcal{D}_{sr}$.

Moreover, in this case, $F_{rd}$ can be re-transmitted from the RN to DN within the time during which $F_{sr}$ is being received at RN. Note that in this case, $F_{rd}$ could not be scheduled in such a way that it might complete its re-transmission before $F_{sr}$ is entirely received. As a result, the best suitable $t_{ir}$, for this case, is somewhere during $F_{sr}$ reception so that the re-transmission of $F_{rd}$ ends right after the reception of the last bit in $F_{sr}$.

$\mathcal{D}_{rd}$ could not be this much higher so that the buffer is subject to overflow. If the BOC is satisfied, the RN selects that channel for re-transmitting $F_{rd}$ at its corresponding rate, $\mathcal{D}_{rd}$. On the other hand, if the BOC is not satisfied, the algorithm checks BOC for the second-best SNR channel(s) available or it adjusts the rate at the same previously chosen channel. Whichever situation satisfies BOC, that specific SNR channel is chosen to compute the packet out schedule time:

$$t_{or}^{HS} = t_{ir} + T_{sr} - T_{rd}. \tag{3.3}$$

$$T_{tot}^{HS} = t_{ir} + T_{sr} \tag{3.4}$$

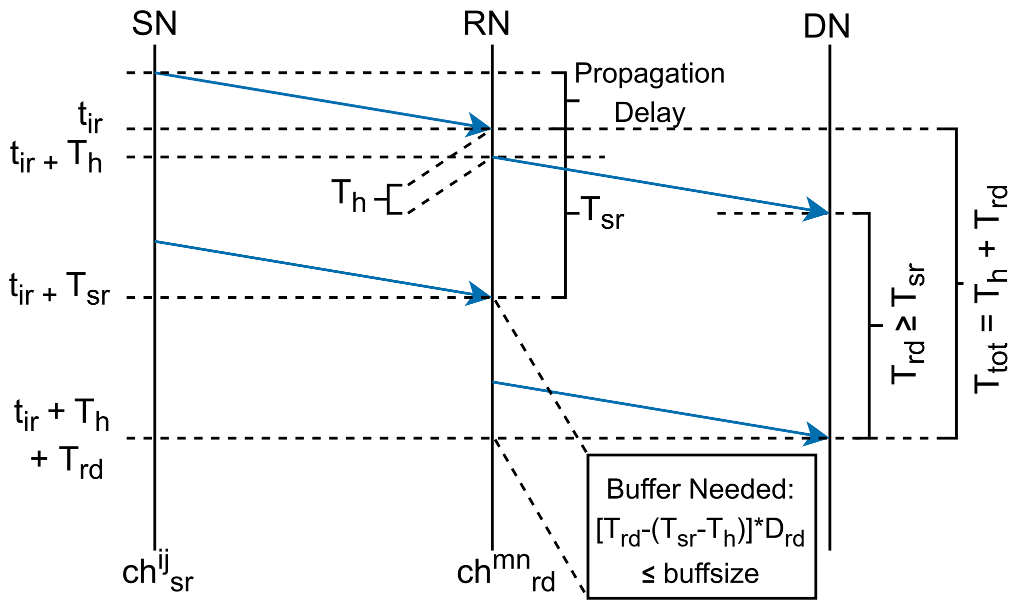



Figure 3.6: Packet timing diagram for the case where only lower SNR channels are available.

#### CASE 2 - Only lower SNR channel(s) available

A detailed timing diagram for the case of a lower SNR channel is shown in Fig. 3.6. In this case, the RN has lower SNR channel(s) available than $\text{SNR}_{sr}$ for $F_{rd}$ re-transmission. This case is not a preferable one, as it will result in the longer value of $T_{tot}$ when compared to $T_{tot}^{HS}$ as illustrated in Fig. 3.4. However, it would still be less than the total time taken by the DF method. Since the RN has only lower SNR channels available than $\text{SNR}_{sr}$, it picks the channel with the maximum SNR among the available channel list and checks the BOC:

$$(T_{rd} - (T_{sr} - T_h))\, \mathcal{D}_{rd} < \text{ buffer-size}. \tag{3.5}$$

Here, $T_{rd}$ is the time span of $F_{rd}$ if it is to be sent on this picked SNR with its corresponding rate [15], which would be lower than $\mathcal{D}_{sr}$ for this case. BOC checks if the RN has enough available buffer for storing $(T_{rd} - T_{rd})\mathcal{D}_{sr}$ number of bits of $F_{rd}$ packets after the reception of $F_{sr}$ completes. In this case, $F_{rd}$ can be sent at a much lower rate, $\mathcal{D}_{rd} < \mathcal{D}_{sr}$; however, it will result in requiring a larger buffer-size. Note that in this case, $F_{rd}$ should be scheduled anytime after reading the header of $F_{sr}$ so that the BOC is satisfied. The best suitable $t_{ir}$ for this case is at $(t_{ir} + T_h)$.

If the BOC is satisfied, the RN selects that channel for re-transmitting $F_{rd}$ at its corresponding rate, $\mathcal{D}_{rd}$. Otherwise, the algorithm accesses the updated SNR$_{list}$ in the database with the recent spectrum sensing results to find any better SNR channel or any lower SNR channel, satisfying the BOC. For the case satisfying BOC, the corresponding SNR channel is chosen to compute the packet out schedule time:

$$t_{or}^{LS} = t_{ir} + T_h. \tag{3.6}$$

$$T_{tot}^{LS} = t_{ir} + T_h + T_{rd} \tag{3.7}$$

### 3.4.4 Computational Complexity

The computational complexity of proposed heuristic can be noticed from the pseudo-code provided in Algorithm 1. This algorithm has three WHILE loops. The outer WHILE

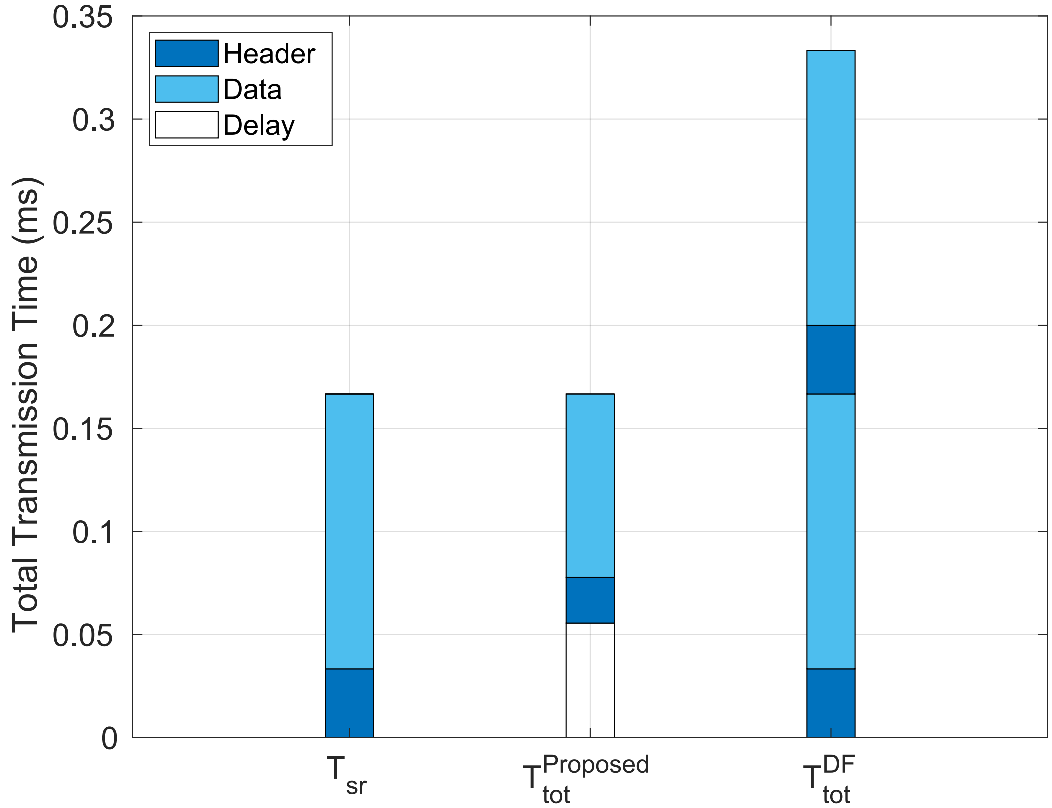



Figure 3.7: High SNR channel case, with one interconnecting RN.

loop always runs where as, out of the inner two WHILE loops, only one of the inner WHILE loop runs being in and IF ELSE condition. Hence the algorithm runs as a two nested WHILE loops and its computational complexity can be given as $O(n^2)$.

## 3.5 Performance Evaluation

Proposed algorithm is evaluated using simulations constructed in MATLAB. The simulation results are presented in this section.

Fig. 3.7 depicts the case of one SN sending data frame to DN via one RN. Here, F = 1000 bits, $\mathcal{D}_{sr}$ = 6 Mbps, $\mathrm{SNR_{sr}} \leq 0$, buffsize=500 bits, number of free channels = 50 having SNR range of 12 dB to 20 dB. The $T_{sr}$ bar, in the figure, shows the time span of the SN to RN data frame. This is the case where at the reception of data frame, the RN finds a higher SNR channel available in its spectrum sensing channel database. It can be noticed from the figure that in the presence of higher SNR free channels, the data frame received at the RN can be forwarded to the DN at a higher rate (i.e., the $T_{tot}^{Proposed}$ bar), as compared to the rate at which it was received from SN, hence reducing the total transmission time. On the other hand, the time taken by the conventional DF method is twice as much as that of $T_{sr}$. This is because of waiting for the same channel to become free for re-transmission.

Next, a simulation of SN to DN data frame transmission via five RNs is conducted. Here, F = 1000 bits, $\mathcal{D}_{sr}$ = 6 Mbps, $\mathrm{SNR_{sr}} \leq 0$, buffsize=500 bits, number of free channels = 50 having SNR range of -5 dB to 9 dB. The result of this simulation is reported in Fig. 3.8. If all the five RNs have an even higher SNR channel available (than the SNR at which the packet was initially received at the previous RN), then the packet is forwarded to DN in a significantly smaller time using multiband communication.

On the contrary, the time taken by the conventional DF method is fives times of $T_{sr}$, due to waiting for the channel to get free at each RN, and re-transmitting on the same channel.

CASE 2 is depicted in Fig. 3.9, with a single RN. Here, F = 1000 bits, $\mathcal{D}_{sr}$ = 9 Mbps having SNR range of 0 dB to 3 dB, buffsize=1000 bits, number of free channels = 50 having SNR range of -5 dB to 0 dB. At the reception of data frame, if the RN does not

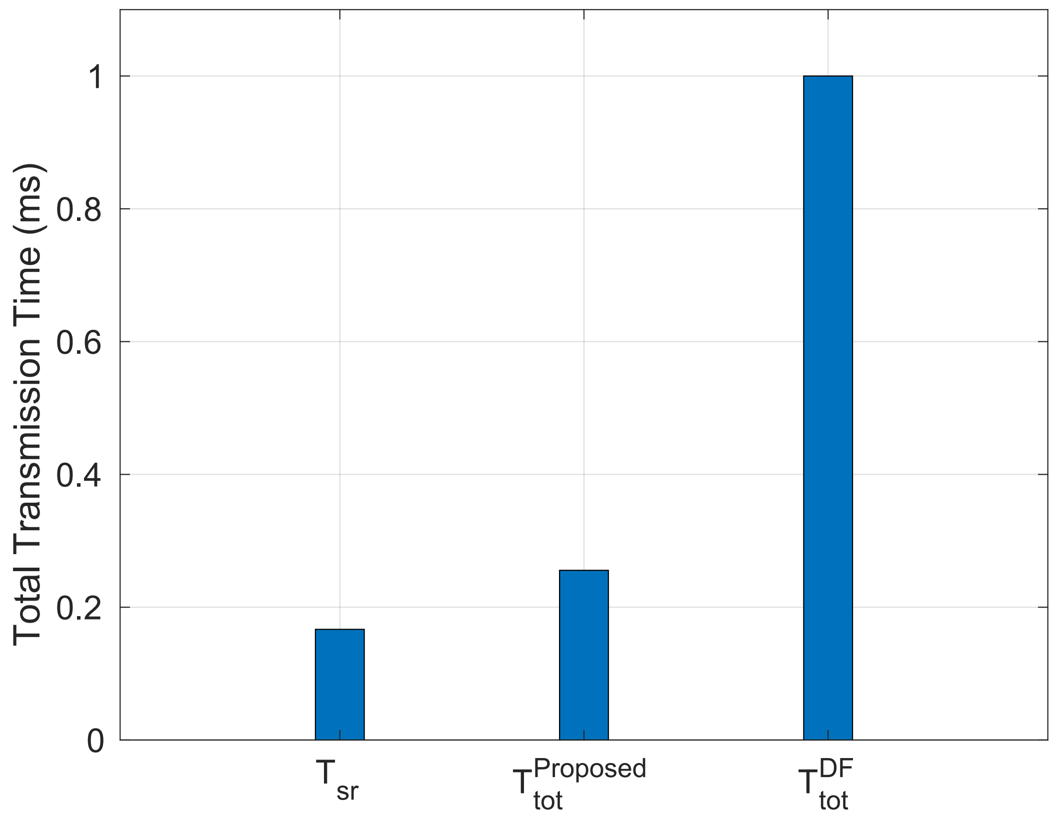



Figure 3.8: High SNR channel case, with five RNs.

have a higher SNR channel available, the RN chooses the channel with the maximum SNR among the available low SNR channel list and starts transmission on that channel using multiband communication.

It can be noticed from Fig. 3.9, that by using a different channel for re-transmission, even if a higher SNR channel is not available, the frame re-transmission could be completed in a much shorter time as compared to the single band DF method. Note that the conventional DF method takes twice as much time as the original frame's time span.

Fig. 3.10 demonstrates the comparison of DF and the proposed algorithm with different

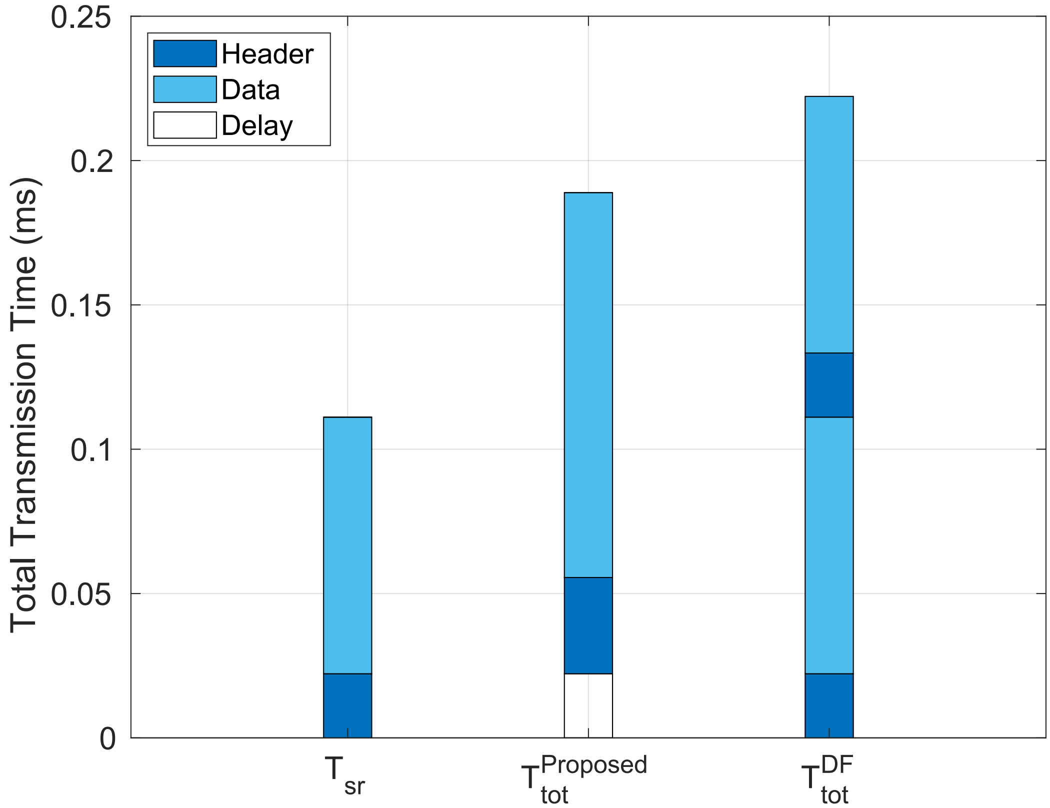



Figure 3.9: Low SNR channel case, with one interconnecting RN.

numbers of RNs. It can be seen from the figure that as the number of RNs increases, the time taken by the conventional DF method increases as a multiple of the number of RNs. On the other hand, proposed algorithm completes the transmission within a greatly reduced duration. Furthermore, it is worth mentioning that this reduction in the transmission time improves as the number of relays increases, which is a preferable feature for new applications such as the Internet of Things (IoT) in densified small cells of 5G and beyond networks.

Next, Fig. 3.11 demonstrates the shortcoming of using the centralized approach in

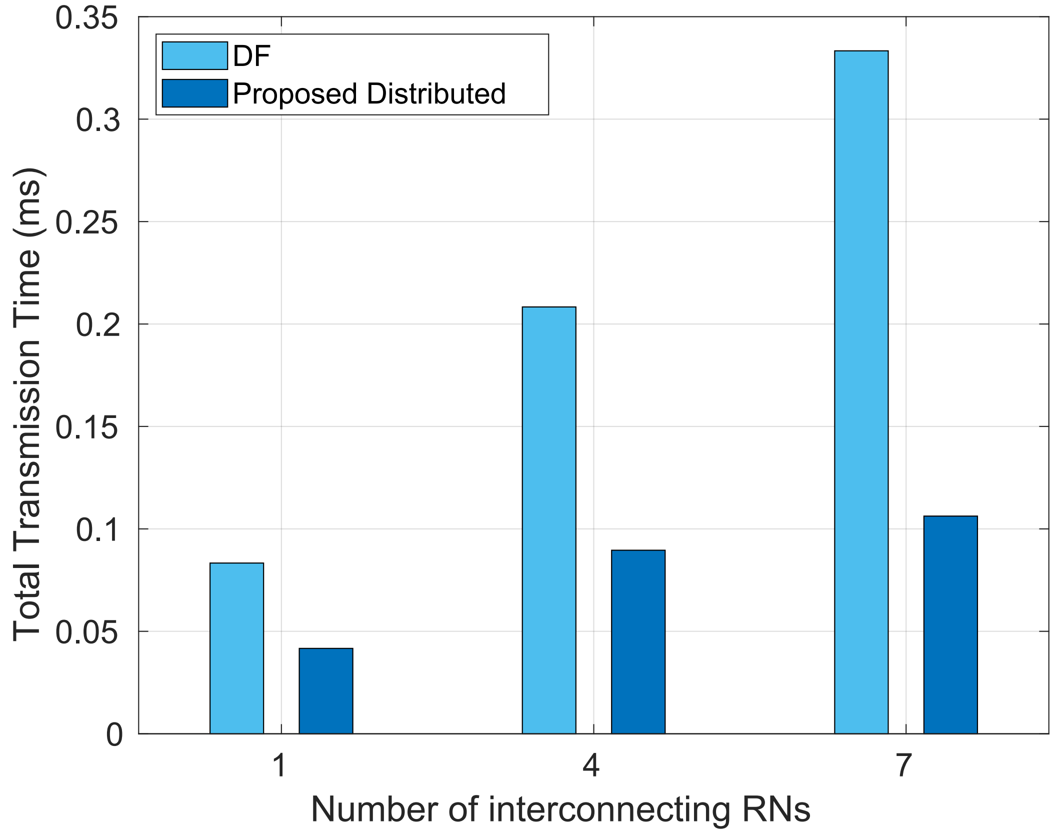



Figure 3.10: Comparison with different number of RNs.

contrast with proposed distributed algorithm. Here, F = 1000 bits, $\mathcal{D}_{sr}$ = 24 Mbps having SNR range of 9 dB to 12 dB, buffsize=800 bits, number of free channels = 50 having SNR range of -5 dB to 18 dB. As the number of RNs increases, the execution time required by the centralized approach increases dramatically. For higher number of RNs, it exceeds the order of seconds violating the milliseconds order delay requirement in 5G communication. On the other hand, proposed heuristic-based distributed approach takes much lower time to complete its execution and reach localized channel allocation decisions.

For example: for the six relays in Fig. 3.11 (F = 1000 bits, $\mathcal{D}_{sr}$ = 24 Mbps having

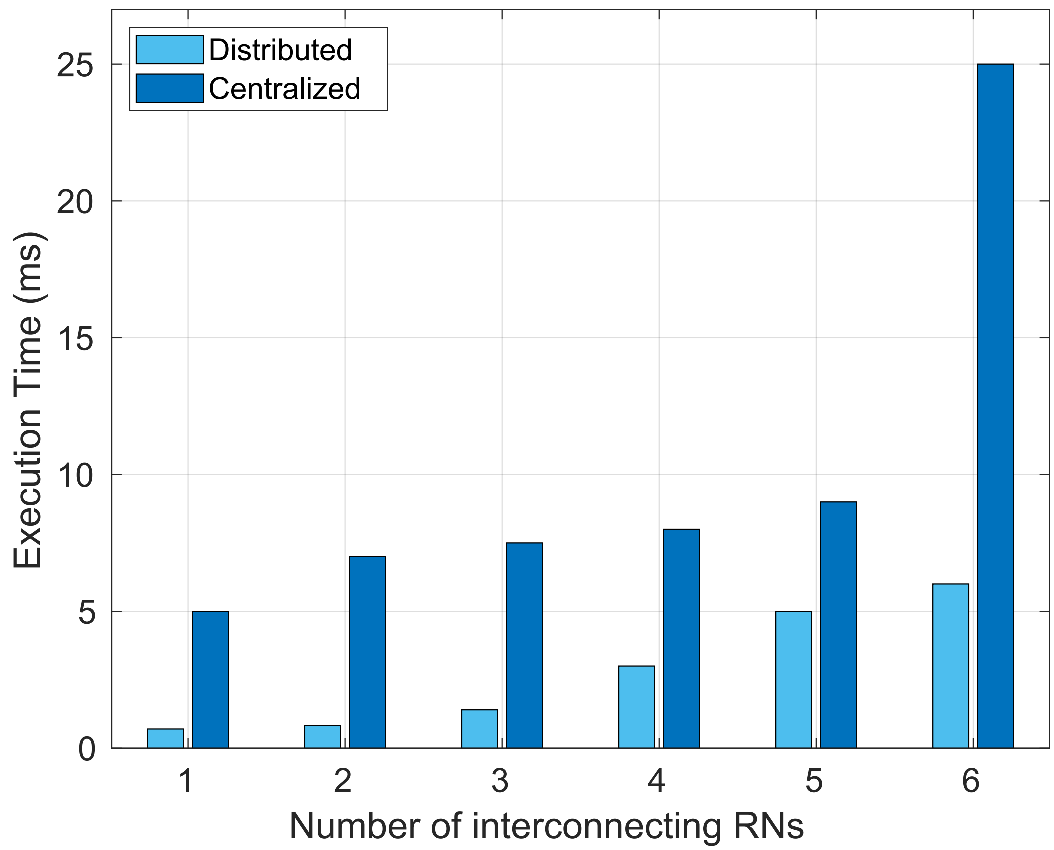



Figure 3.11: Execution time comparison between proposed centralized and distributed approaches.

SNR range of 9 dB to 12 dB, buffsize=1000 bits, number of free channels = 50 having SNR range of -5 dB to 18 dB.), the execution time for proposal is approximately 6ms, whereas for the centralized approach, it is 25ms. In other words, the execution time of the centralized approach is over four times than that of the distributed counterpart, and it even worsens as the number of RNs increases. This high complexity of the centralized algorithm renders it ineffective for deployment in practical applications. However, it can be used as a benchmark for proposal and future distributed variants.

## 3.6 Conclusion

In this chapter, the capacity versus coverage trade-off has been taken into account for next generation communication networks. Along which, the shortcoming of existing relay-assisted techniques is also discussed so as to improve the service coverage.

The research gap is identified in the existing literature for using multiple bands in RNs simultaneously without considering the impact of delay. This issue is addresses by formulating an optimal heterogeneous band/channel allocation method for the RNs to achieve minimum latency.

Because the centralized approach is not practical for deployment in the resource-constrained RNs such as device-to-device networks, a heuristic-based distributed channel allocation algorithm is proposed with the local information available at individual RNs.

Conducted simulation results demonstrated that proposed distributed approach outperformed the conventional DF method, and also scaled well with the increasing number of RNs in contrast with the centralized method to solve the formulated optimization problem.

## 3.7 Publications Resulted from This Chapter

- B. Mughal, Z. M. Fadlullah and S. Ikki, "Centralized Versus Heuristic-Based Distributed Channel Allocation to Minimize Packet Transmission Delay for Multiband Relay Networks," *IEEE Networking Letters*, vol. 2, no. 4, pp. 180-184, Dec. 2020.

# Chapter 4

# Allocation Schemes for Relay Communications: A Multi-Band Approach Using Game Theory

**Motivation:** As previously proposed distributed model is based in greedy nature, hence there is still enough room for improvement. In this chapter a centralized oracle is presented, based on a bipartite graph model for the multi-band multi-channel allocation to RNs, and indicate why this is a non-viable for deployment. Therefore, with reference to the previously proposed greedy approach, a better online, distributed multi-band/channel allocation strategy is explored by proposing a sequential game-theoretic algorithm. Simulation results demonstrate that proposed game-theoretic approach significantly outperforms the traditional distributed and centralized methods.

## 4.1 Introduction

In this chapter, multi-band-based both centralized and distributed approaches are proposed, to solve the latency issue in relay communication. For centralized approach graph theory is used, modeling the problem as a bipartite graph and solving it using optimization theory tools. In the mathematical field of graph theory, a bipartite graph is a graph, whose vertices can be divided into two disjoint and independent sets. Such that every edge of the graph connects a vertex in first set to one in the other set. The vertex sets are usually called the parts of the graph. In this work the first distinct set is the RNs and the other distinct set is the possible power frequency values that can be used by RNs in order to forward data packet.

For distributed approach a game theory-based model is proposed which, not only solves the problem most efficiently but also fulfils all the related constrains. Game theory is the study of strategic interactions among two or more players also known as decision-makers. This study include the process of modeling the corresponding mathematical structure, on the basis of certain moves, set of rules and outcomes [5].

Depending on the characteristics of the players, strategies, rules and environment, the particular game can be modeled using various forms: cooperative/non-cooperative, symmetric/asymmetric, zero-sum/non-zero-sum, simultaneous/sequential, discrete/continuous games [32]. In this work RNs are modeled as game players and propose a sequential game with perfect information. The proposed game is then represented as a decision tree for further analysis.

The remainder of the chapter is organized as follows. Considered centralized oracle, based on a bipartite graph, is presented and the formal problem of multi-band multi-channel allocation for the RNs is formulated in section 4.2.

Next, in section 4.3, a sequential game theoretic algorithm is presented as an advanced distributed channel allocation method. Simulation results are reported in section 5.6 to compare the performances of the centralized oracle and proposed distributed techniques. Concluding remarks are presented in section 6.6.

## 4.2 Considered Centralized System Model and Problem Description

This section presents considered centralized system model, based on a balanced bipartite graph and then constructs an optimization problem formulation. As shown in Fig. 1.2, the system model assumes one CN, acting as an oracle and having a complete information of the network, i.e., spectrum sensing results [11]; SNR estimation of free channels; buffer size; available power; distance between interconnecting SN, RN/s and DN; and so forth. In addition, the CN is responsible for making decisions and allocating resources, such as power and frequency.

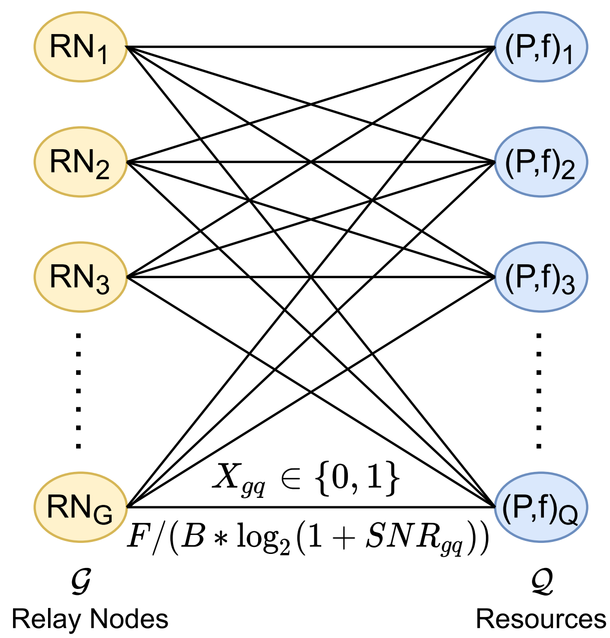



Figure 4.1: Bipartite Graph for Central Oracle.

### 4.2.1 Bipartite Graph-based Oracle

The relay-based network can be represented as a combinatorial, constrained-weighted bipartite graph matching problem $\mathcal{C} = \{\mathcal{G}, \mathcal{Q}, \mathcal{E}\}$ consisting of two disassociated sets of vertices $\mathcal{G}$, $\mathcal{Q}$, and a set of edges $\mathcal{E} = \mathcal{G} \times \mathcal{Q}$, as depicted in Fig. 4.1. An edge $\mathrm{X_{gq}}$ connects a vertex $RN_g \in \mathcal{G} = \{RN_1, RN_2, ..., RN_{\mathcal{G}}\}$ with a vertex $(P, f)_q \in \mathcal{Q} = \{(P, f)_1, (P, f)_2, ...,$ $(P, f)_{\mathcal{Q}}\}$ and has an associated weight $\mathrm{U}gq$, where $(P, f)$ represents a pair of power level and available channel frequency. The objective is to find a matching $\mathrm{X} \subseteq \mathcal{E}$ that associates every vertex in $\mathcal{G}$ with a vertex in $\mathcal{Q}$.

In a bipartite graph, when the cardinalities of the vertex sets are equal (i.e., $|\mathcal{G}| = |\mathcal{Q}|$), perfect matching can be attained [3]. The vertices $\mathrm{RN_g}$ represent the RNs that belong to the same set $\mathcal{G}$, whereas the vertices $(P, f)_q$ represent the allotable resources, which are denoted by set $\mathcal{Q}$. In this work, the packet time span $\mathbf{U}_{g,q}$ is considered to be the weight of the edge connecting $g^{th}$ RN to the $q^{th}$ $(P, f)$ pair. Here, $\mathbf{U}_{g,q} = \mathrm{F}/(\mathrm{B}\log 2(1+\mathrm{SNR_{gq}}))$, F is the packet size in bits, B is the bandwidth of the resource and $\mathrm{SNR_{qg}}$ is the function of the $q^{th}$ $(P, f)$ pair connected to the $g^{th}$ RN. The goal is to choose the edges/connections that grant minimum weights, i.e., to find $\mathrm{X_{gq}} \in \{0, 1\}$, where 0 means discarding the connection while 1 indicates keeping the connection.

### 4.2.2 Problem Formulation

To describe considered problem, let $\mathcal{G}$ contain the actual number of RNs plus dummy RNs in order to construct a balanced bipartite problem [3]. Also, let $\mathcal{Q}$ denote the number of power level pairs and available free channels. Let $\mathbf{U}$, $\mathbf{T}$, $\mathbf{SNR}$, and $\mathbf{BO}$ be $\mathcal{G} \times \mathcal{Q}$ matrices. Here, $\mathbf{U} = \mathbf{T}$, containing the outcomes of all the edges and is equal to the time span of the forwarded data packet against each power frequency pair. $\mathbf{BO}$ is the buffer overflow matrix against each edge, and $\mathbf{SNR}$ denotes the corresponding SNR for each pair.

Let $\mathbf{P}$ and $\mathbf{f}$ be the row vectors of size $\mathcal{Q}$, representing the corresponding power values and available channel frequencies for all the pairs, respectively. Now, considered problem can be formally modeled as an optimization problem as follows:

$$\min \quad \sum_{g=1}^{\mathcal{G}} \sum_{q=1}^{\mathcal{Q}} \mathbf{X}_{g,q} \cdot \mathbf{U}_{g,q}, \qquad\qquad s.t. \tag{4.1a}$$

$$\sum_{g=1}^{\mathcal{G}} \sum_{q=1}^{\mathcal{Q}} \mathbf{X}_{g,q} \cdot \mathbf{BO}_{g,q} = 0, \tag{4.1b}$$

$$\sum_{q=1}^{\mathcal{Q}} \mathbf{X}_{g,q} \cdot \mathbf{P}_{q} \leq P_{avail}, \quad \forall\, g, \tag{4.1c}$$

$$\mathbf{X}_{g,q} + \mathbf{X}_{g+1,q} \leq 1, \quad \forall\, q, g, \tag{4.1d}$$

$$\sum_{q=1}^{\mathcal{Q}} \mathbf{X}_{g,q} \cdot \mathbf{SNR}_{\mathrm{g,q}} \geq \mathrm{SNR}_{\mathrm{min}}, \quad \forall\, \mathrm{g}, \tag{4.1e}$$

$$\sum_{q=1}^{Q} \mathbf{X}_{g,q} = 1 \quad \forall\, g, \tag{4.1f}$$

$$\mathbf{X}_{g,q} \in \{0, 1\}, \tag{4.1g}$$

where, $\mathbf{X}_{g,q}$ is a $\mathcal{G} \times \mathcal{Q}$ matrix, binary decision variables which are unknown. The details of above problem are given below:

1. Objective function (4.1a) aims to minimize the sum of all forwarded packets' time spans.

2. The equality constraint (4.1b) checks if all the RNs have a buffer overflow equal to zero.

3. The inequality constraint (4.1c) aims to keep the power usage in accordance with the available power of the RNs.

4. On the other hand, the inequality constraint (4.1d) forces the system to use a multi-band/channel setting.

5. Constraint (4.1e) maintains the QoS of each forwarded data packet.

6. The equality constraint (4.1f) compels each RN to pick exactly a single power frequency pair to keep forwarding the data packet.

7. Constraint (4.1g) is for characterizing the unknown matrix $\mathbf{X}$ as a Boolean variable.

### 4.2.3 Oracle-based multi-band multi-channel allocation to relays

The steps of the aforementioned centralized bipartite graph-based oracle for an optimal multi-band/channel allocation to the RNs are enumerated in Algorithm 2 and are explained here, in order to minimize the objective function (4.1a):

1. First, the CN accesses its central spectrum sensing results and generates $(P, f)$ pairs with all possible combinations of power levels and available channel frequencies.

2. After generating these pairs, the CN collects information on the packet size F directly from the SN (using any information channel).

3. Then, CN collects the information on distances between each RN and the power available for each.

4. Next, the CN generates matrices of time span, buffer overflow, and SNR against each power frequency pair for each RN.

5. After this, the CN generates the equality and inequality matrices for an optimization solver.

6. The solver then returns the solution matrix to the CN, which then gets forwarded to each RN.

7. Along with the solution to the problem, the CN now has the scheduling times for the packets of each RN and shares that information with all RNs.

8. The SN commences transmitting and each RN, at this point, already knows which power frequency is to be selected in order to forward the received data packet to the DN.

While this centralized optimization algorithm requires an oracle (i.e., a full knowledge on the system), it is not progressive with respect to the individual RNs. Because it is not possible to incorporate the ability to the optimization problem to assess local buffer overflow conditions at the RNs as it depends on the previous nodes' decisions regarding data rate. In the centralized model, all decisions are taken at once by the central node, and therefore, it can only check the buffer overflow condition for all the nodes at once.

With an increasing number of RNs, however, this approach takes non-deterministic time and adds exponential latency to the system. This type of approach would only be suitable for devices with limited power, where it is capable of finding the global optimal solution for the network with infinite buffer sizes. Due to such impracticality, a distributed solution is desired to solve the aforementioned problem.

**Algorithm 2** Centralized oracle-based multi-band multi-channel selection.

**Input:** **U**, **SNR**, **BO**, **P** (in addition to Algorithm 2 inputs)

**Output:** $t_{or,g}$, **X**

1: The CN accesses its central spectrum-sensing results and generates $(P, f)$ pairs;

2: SN sends information on packet size F to CN.

3: CN collects information on $P_{g,avail}$ and $d_{g,(g+1)}$ for each RN;

4: CN generates matrices **U**, **SNR**, and **BO**;

5: CN generates the equality and inequality matrices for the solver;

6: The solver returns the solution **X** by solving problem 4.1;

```
t_{ir,g} = 0;
for g=1:G do
    T_{g,(g+1)} = U(g,:) · X(g,:)
    if T_{(g-1),g} > T_{g,(g+1)} then
        t_{or,g} = t_{ir,g} + (T_{g,(g-1)} - T_{g,(g+1)});
    else
        t_{or,g} = t_{ir,g} + T_{h,(g-1)};
    t_{ir,g} = t_{or,g};
    T_{(g-1),g} = T_{g,(g+1)};
    return t_{or,g}
```

## 4.3 Proposed Game Theoretic Distributed Multi-Band Multi-Channel Selection Algorithm

Previously in [33], a time-sensitive multi-band relay communication model is proposed, using a distributed greedy algorithm. Customized greedy heuristic can be regarded as a reference solution as the first intuitive solution to solve the optimization problem in a distributed manner. However, due to its greedy nature, there is a room for further improving the quality of the solutions, which is investigated using a sequential game theoretic approach in this section.

### 4.3.1 Sequential Game Model: Motivation and Preliminaries

A sequential game involves a model in which one player performs an action (distributively) before other players make their decisions (i.e., no two players make moves at the same time) [34]. In considered case, a particular RN can start transmitting at a particular power and channel (makes a decision) only after it starts receiving the packet from the previous node.

Importantly, the later players must have some information of the earlier ones' choices; otherwise the difference in time would have no strategic effect. In considered case, the packet header received from the previous node provides this information to the particular RN.

Perfect information is often confused and used interchangeably with complete information; however, here an important distinction is made. In this model, it is assumed that RNs receive perfect information in the packet header from the prior node. In proposed work, complete information is not used, i.e., all RNs having information regarding the entire network as considered with the centralized oracle. Hence, we construct a sequential combinatorial game with perfect information, not with complete information, using the following extensive form representation [35] and entities:

1. Players: In considered case, RNs are regarded as players.

2. The information available to each player: Here, RNs hold information about the packet size/length and the rate at which data is received.

3. Actions choices for each player at point of decision: Actions available to every RN include selecting a power value pair out of $\mathcal{P}$ possible power values $P = [P_1, P_2, ..., P_{\mathcal{P}}]$ and $\mathcal{F}$ available channels $f = [f_1, f_2, ..., f_{\mathcal{F}}]$, at which the data can be forwarded to the next player.

4. The payoffs for each outcome: Here, the payoff is a Boolean representation. Either all the constraints are met or not against each power and frequency channel pair.

In game theory, these elements are typically used, along with strategies available for each player. Here, each RN's strategy is to choose a power frequency pair that satisfies the utility function and fulfills all constraints. The utility function of the modeled game is to reduce the forwarded packet's time span. The sequential aspects, characterized by the extensive form representation can also be depicted as a decision tree given in Fig. 4.2. The decision tree demonstrates the possible ways of playing th considered game. $D_{sr1}$ denotes the initial data rate chosen by the source/SN in accordance with its own channel condition. To trigger the relay communication, this value is considered the starting/initial point for the necessary algorithm to be designed.

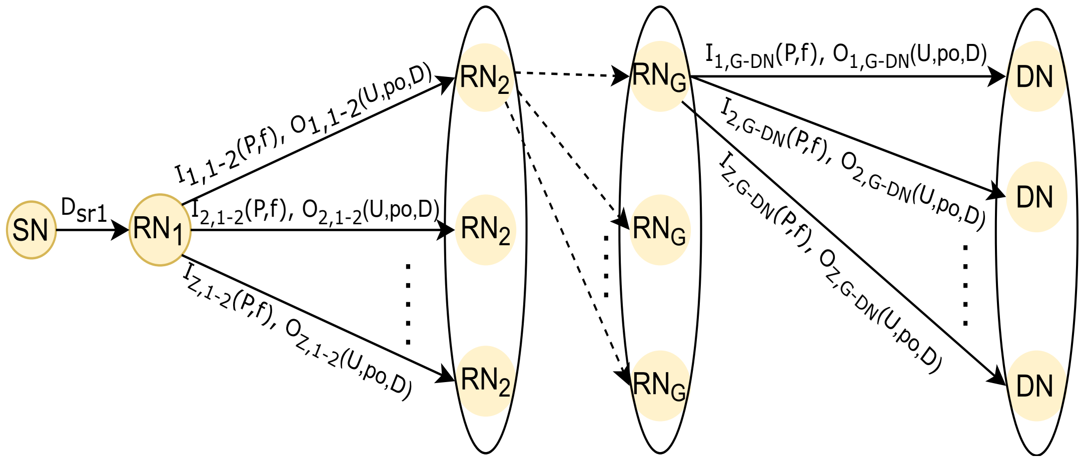



Figure 4.2: Decision tree of proposed sequential game model for RNs.

### 4.3.2 Reformulation of the Original Problem

Now, the original optimization problem in Eq. (4.1a) is transformed into the distributed game model for $g^{th}$ player (RN) as follows:

$$\mathrm{U} = \min\{T_{g,(g+1)}(P_{g,(g+1)}, f_{g,(g+1)})\}, \quad s.t.(po \in \{0,1\}) \tag{4.2a}$$

$$BO(\mathcal{D}_{(g-1)}, \mathrm{buffsize_g}) = 0, \tag{4.2b}$$

$$f_{(g-1),g} \neq f_{g,(g+1)}, \tag{4.2c}$$

$$\mathrm{SNR_{g,(g+1)}} \geq \mathrm{SNR_{min}}, \tag{4.2d}$$

$$0 \leq P_{g,(g+1)} \leq P_{g,avail}, \tag{4.2e}$$

where,

$$T_{g,(g+1)} = \frac{\mathrm{F}}{\mathcal{D}_{g,(g+1)}} \tag{4.3}$$

$$\mathcal{D}_{g,(g+1)} = B\log_2(1 + \mathrm{SNR_{g,(g+1)}}) \tag{4.4}$$

$$\mathrm{SNR_{g,(g+1)}} = \frac{\mathrm{P_{g,(g+1)}h_{g,(g+1)}}}{\mathrm{noise}} \tag{4.5}$$

$$h_{g,(g+1)} = \frac{c}{4\pi f_{g,(g+1)}d_{g,(g+1)}}, \tag{4.6}$$

and $P_{g,(g+1)}$, $f_{g,(g+1)}$, $T_{g,(g+1)}$ are the chosen power, channel, data time span respectively. These are used to transmit data packet from the $g^{th}$ RN to the next $(g+1)^{th}$ node. The details are given below:

1. The constraint (4.2b) states that while using the particular $(P, f)$ pair, the buffer should not overflow, and this is checked against the $(g-1)^{th}$ RN's data rate $(\mathcal{D}_{(g-1)})$ and the $g^{th}$ RN's available buffer size (buffsize).
2. Constraint (4.2c) is defined as forcing the player to use the multi-channel/multiband opportunity in order to reduce latency.
3. Constraint (4.2d) is the respective QoS constraint to maintain the data quality.
4. Constraint (4.2e), however, is the power limited constraint of movable devices.
5. *po* is the payoff Boolean variable representing 1 for all constraints met and 0 otherwise.

**Algorithm 3** DISTRIBUTED SEQUENTIAL GAME.

**Input:** $P_{g,avail}$, buffsize, $\text{SNR}_{\min}$, $d_{g,(g+1)}$

**Output:** $t_{or,g}$

1: Set $t_{ir,g} = t_{or,(g-1)}$ at the reception of packet from the prior player;

2: Read received packet header;

3: Collect/learn information on $\mathcal{D}_{(g-1),g}$, F, $f_{(g-1),g}$;

4: Calculate $T_{(g-1),g} = F/(\mathcal{D}_{(g-1),g})$;

5: Generate array $\mathcal{F}$ from local spectrum-sensing results;

6: Generate array $\mathcal{P}$ with all possible power levels;

7: Generate $(P, f)$ with all possible combinations of set $\mathcal{P}$ and $\mathcal{F}$;

8: Call Procedure (see Algorithm: 4);

**if** $T_{(g-1),g} > T_{g,(g+1)}$ **then**
  $t_{or,g} = t_{ir,g} + (T_{g,(g-1)} - T_{g,(g+1)})$;
**else**
  $t_{or,g} = t_{ir,g} + T_{h,(g-1)}$;
**return** $t_{or,g}$

### 4.3.3 Envisioned Game Theoretic Solution

Now, Algorithms 3 and 4 are proposed so as to progressively minimize the utility function Eq. (4.3) at the RNs. The steps of the algorithm are explained below.

1. The SN transmits data of size F (bits) to the $\text{RN}_1$.

2. The player $\text{RN}_1$ extracts the perfect information, including the incoming data's size, F and rate, $\mathcal{D}$ from the header of the received packet.

3. The player $\text{RN}_1$ picks all the power value and channel $(P, f)$ pairs out of the $\mathcal{P}$ possible

power values $P = [P_1, P_2, ..., P_{\mathcal{P}}]$ and $\mathcal{F}$ available channels $f = [f_1, f_2, ..., f_{\mathcal{F}}]$, and calculates the corresponding outcome of the utility function (U) as well as the payoff ($po$) for each pair.

4. The utility function U is the corresponding time span, and T of each packet is to be forwarded on the basis of a particular $(P, f)$. The payoff, $po$ is a Boolean variable indicating 1 for the successful fulfillment of all constraints and 0 otherwise.

5. The player $\mathrm{RN}_1$ then selects one pair out of all others, resulting in $po = 1$ and a minimum U among them. Therefore, it selects the pair with the minimum U, which fulfills all the constraints.

6. Since this is a sequential game, each player (RN) takes turns in a linear, progressive fashion. When the first RN finalizes its selection of $(P, f)$, it can forward packets to the next node, providing information about F and $\mathcal{D}$ in the header. Now, the next node is ready to play the game as was done by the $\mathrm{RN}_1$.

The proposed algorithm is fast and simple for a limited number of branches. This approach makes decisions distributively and on the basis of current/updated local band/channel conditions, hence capable of producing more reliable results. Moreover, all constraints can be handled by this approach, rendering proposed game theoretic algorithm suitable for both limited power and buffer size.

**Algorithm 4** PROCEDURE FOR COMPUTING THE UTILITY $U$.

**Input:** $(P, f)$, $T_{(g-1),g}$, $f_{(g-1),g}$, $t_{ir}$

**Output:** $U = T_{g,(g+1)}$, $\mathcal{D}_{g,(g+1)}$, $(P, f)_{g,(g+1)}$

**for** $q = 1 : size(P, f)$ **do**

1: BC = PC = SC = FC = zeros(1,Q);

2: Calculate $T(q)$, $\mathcal{D}(q)$, and SNR$(q)$ using Eqs. (4.3), (4.4), and (4.5), respectively;

3: Check buffer overflow constraint using Eq. (6.4b):

**if** $T(q) < T_{(g-1),g}$ **then**
  **if** $(T_{(g-1),g} - T(q)) * \mathcal{D}_{(g-1),g} \leq buffsize$ **then**
    $BC(q) = 1$;
**else**
  **if** $(T(q) - T_{(g-1),g}) * \mathcal{D}(q) \leq buffsize$ **then**
    $BC(q) = 1$;

4: Check multiband/channel constraint using Eq. (**??**):

**if** $f(q) \neq f_{(g-1),g}$ **then**
  $FC(q) = 1$;

5: Check QoS constraint using Eq. (6.4c):

**if** $SNR(q) > SNR_{min}$ **then**
  $SC(q) = 1$;

6: Check limited power constraint using Eq. (6.4d):

**if** $P(q) > P_{g,avail}$ **then**
  $PC(q) = 1$;
**if** $BC(q) == 1\ \&\ FC(q) == 1\ \&\ SC(q) == 1\ \&\ PC(q) == 1$ **then**
  $po(q) = 1$;

$indices = find(po == 1)$;
$T_{g,(g+1)} = min(T(indices))$;
$index = find(T == T_{g,(g+1)})$;
$P_{g,(g+1)} = P(index)$;
$f_{g,(g+1)} = f(index)$;
SNR$_{g,(g+1)}$ = SNR$(index)$;
**return** $T_{g,(g+1)}$

### 4.3.4 Computational Complexity

It can be noticed from the pseudo-codes of proposed sequential game approach that Algorithm 4 is being called from Algorithm 3 at step 8. In Algorithm 4, there is only one FOR loop that runs for input of $size(P, f)$, which represents the number of considered power frequency pairs. Hence the computational complexity of proposed sequential game can be given as $O(n)$, where $n = size(P, f)$.

## 4.4 Performance Evaluation

In this section, the performance of proposed game theoretic approach, based on computer-based simulations, is evaluated and compared with the two reference methods (i.e., centralized oracle and distributed greedy heuristic). MATLAB is used to construct simulations with the simulation and environmental parameters used in the earlier work [33] unless otherwise stated.

In Fig. 4.3 (F = 3000 bits, $\text{SNR}_{\text{min}} = -5$ dB, buffsize = 100, $50 \leq d_g(m) \leq 1000$, $0.0001 \leq P_g(W) \leq 0.5$, $0.05 \leq P_{g,avail} \leq 0.5$, $\mathcal{D}_0 = 198$ Mbps), it can be seen that as the number of RNs/players increases, the proposed algorithm has the smallest latency when forwarding the data packet from the SN to the DN. As expected, the greedy approach has a higher latency. When compared with the game and greedy models, the oracle-based method experiences an exponential rise in the packet latency as the number of RNs grows significantly.

Next, in Fig. 4.4 (F = 3000 bits, $\text{SNR}_{\text{min}} = -5$ dB, buffsize = $100, 200, 600$, $50 \leq$

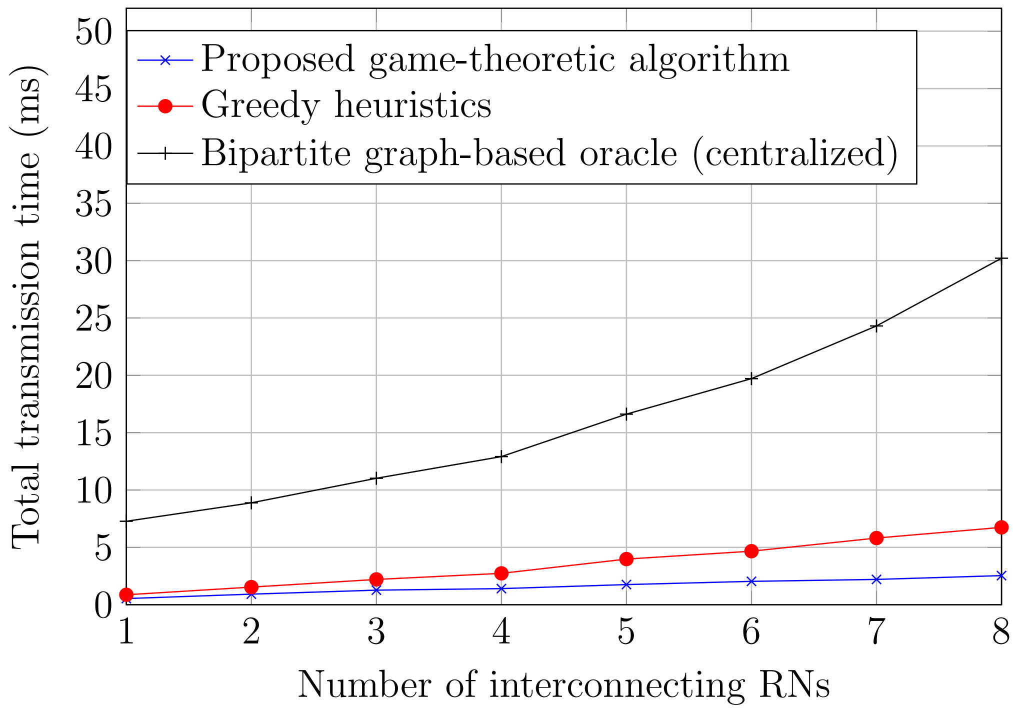



Figure 4.3: Comparison for the total transmission time for the considered methods with a varying number of RNs.

$d_g(m) \leq 100$, $0.01 \leq P_g(W) \leq 3$, $1.5 \leq P_{g,avail} \leq 3$, $\mathcal{D}_0 = 50$ Mbps), three different buffer sizes are considered for the results. It can be seen from the figure that as the available buffer size increases, the number of overflown bits decrease and vice versa. Moreover, in the worst case scenarios, the greedy approach results in the maximum buffer overflow when compared with the proposed game and oracle-based approaches.

Because the greedy approach ignores the power usage consideration, it rsults in depleting all the power available for that particular transmission, as evident from the results in Fig. 4.5 (F = 3000 bits, $\text{SNR}_{\min} = -5$ dB, buffsize = 100, $50 \leq d_g(m) \leq 100$, $0.01 \leq P_g(W) \leq 3$, $2 \leq P_{g,avail} \leq 3$, $\mathcal{D}_0 = 50$ Mbps). When solving the problem cen-

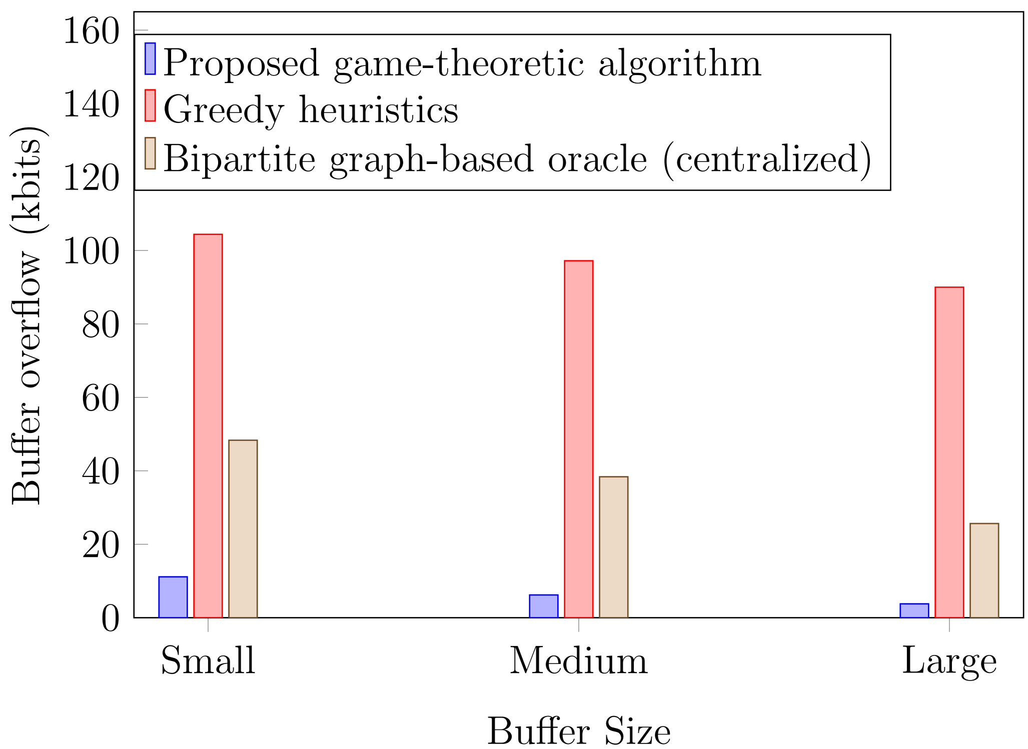



Figure 4.4: Buffer overflow comparison with different buffer sizes for 8 RNs.

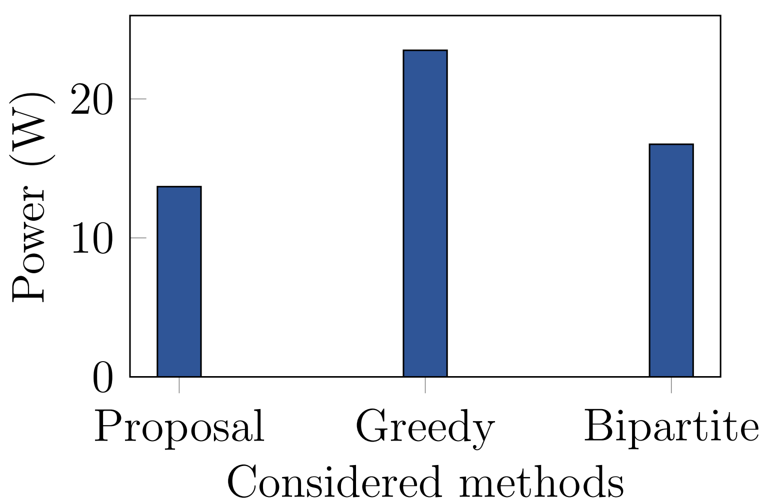



Figure 4.5: Power usage comparison between greedy, game and bipartite approaches for 8 RNs.

trally, the oracle-based approach considers the buffer size to be the maximum, whereas the game theoretic, distributed approach remains tightly constrained with the buffer overflow condition. As a consequence, the proposal uses the least amount of power among all the methods.

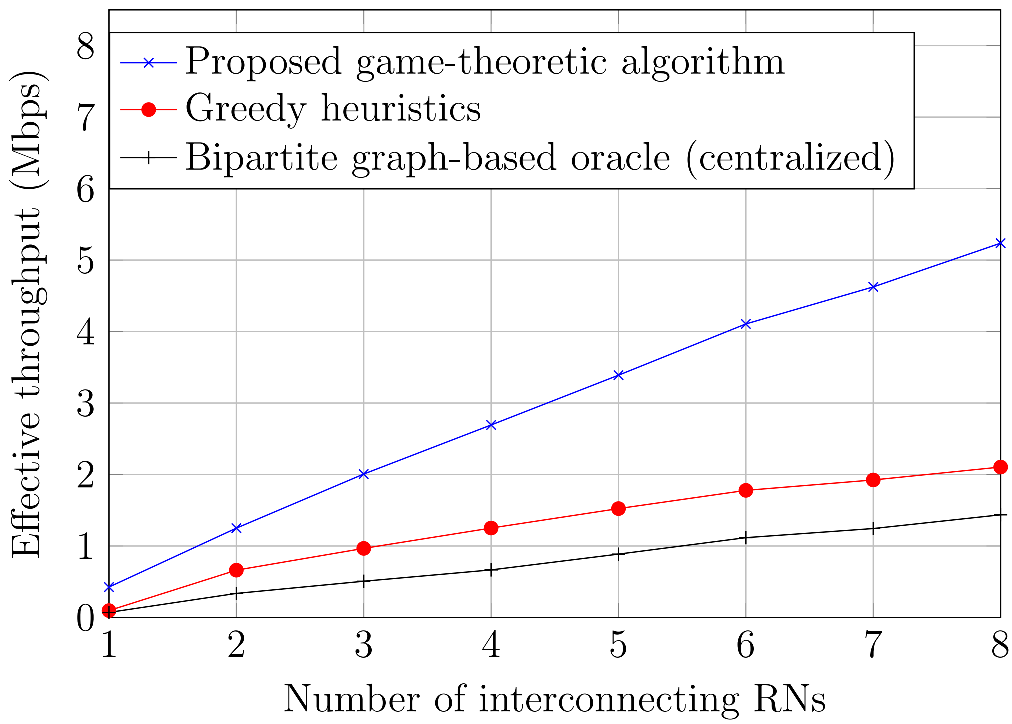



Figure 4.6: Effective throughput comparison with different number of RNs.

The effective throughput is defined as the number of successful packets received during the total time of relaying. It can be seen from Fig. 4.6 (F = 3000 bits, $\mathrm{SNR}_{\min} = -5$ dB, buffsize = 1000, $50 \leq d_g(m) \leq 100$, $0.001 \leq P_g(W) \leq 0.5$, $1 \leq P_{g,avail} \leq 2$, $\mathcal{D}_0 = 1$ Mbps) that the effective throughput of the proposed game theoretic method is the highest, yielding both the most time-efficient and least buffer overflow performances. On the other hand, the oracle-based method has the least effective throughput, demonstrating an exponential rise in the packet latency.

## 4.5 Conclusion

To conclude, in this chapter, the need to optimally assign channels from heterogeneous frequency bands is considered so as to improve the overall transmission time and effective throughput while reducing the energy consumption of the RNs. Due to the practical limitation of a centralized oracle, a sequential game algorithm is explored. Simulation results demonstrated that the performance of proposed model eclipses that of traditionally used approaches including a centralized oracle and a greedy heuristic-based distributed method. In next chapter the detailed complexity analysis of proposed sequential game is discussed.

## 4.6 Publications Resulted from This Chapter

- B. Mughal, Z. Fadlullah, M. M. Fouda and S. Ikki, "Allocation Schemes for Relay Communications: A Multi-Band Multi-Channel Approach Using Game Theory," *IEEE Sensors Letters*, Accepted, 2022.

# Chapter 5

# Nash Equilibrium

**Motivation:** As, in the optimization theory the concept of optimality exists, where the determination is to find the best and top solution from all the feasible solutions. Similarly in game theory the concept of Nash Equilibrium exist, where all the agents can reach to a combined satisfactory solution. This is the motivation behind presenting the Nash Equilibrium analysis of our proposed game theoretic model in this chapter.

## 5.1 Introduction

In this chapter, a brief computational and theoretical analysis of the proposed game theoretical model is provided. Here, the idea of equilibrium for extensive games with perfect information is discussed, particular in the context of wireless communication networks.

## 5.2 Nash Equilibrium in Game Theory

In game theory, in order to find the best strategy for each player, the theorem of Nash Equilibrium is used [6]. Nash Equilibrium is an ideology where each device (player) has some or all the information of the corresponding network and there are no other parameter adjustments (moves/strategy) left that can enhance its performance (utility). To choose its next strategy, each player first looks for the network (other devices) information available to it and then opts for the strategy that can enhance its performance (a step forward).

In other words, it can be said that each player chooses a strategy from which it will never deviate (as any other strategy will not enhance its performance) so long as it cannot go a step higher. Nash Equilibrium is also said to be the most "dominant strategy" leading towards better results out of all other possible moves. Nash strategy is not always the most optimal strategy but is actually the best one.

## 5.3 Nash Equilibrium-Sequential Game Specific

For sequential games with perfect information (also known as "extensive games", "strategic games" and "dynamic games"), in order to find the best strategy for each player, subgame perfect equilibrium (SPE) is found using backward induction [36, 37]. This is the process of studying and analyzing the results of a sequential game played in a forward trend: where the decision-making process starts from the first node and ends at the last one. Now all decision makers are fully aware of the actions chosen by the previous and next players.

Backward induction, on the other hand, occurs when the decision-making process starts from the last node and ends at the first one, which is then used to accomplish equilibrium. For backward induction, first, the game is sectioned into sub-games, as per the rules [38].

Then starting from the last player, the best action for last player is chosen. Then the next-to-last player chooses its best move (this time having the knowledge of moves available to the last player) in order to increase the overall utility. This process is explained below using an illustrative example involving a relay network.

## 5.4 Illustrative Example

Here the sub-game played between $(g-1)^{th}$ and $g^{th}$ RNs is considered, while performing backward induction with the aid of Fig. 5.1. As discussed earlier, each branch has its corresponding payoff and utility functions, and it is assumed that initially, during the forward sequential game, the branches with $po = 0$ are already discarded. Starting in a backward fashion from $g^{th}$ RN. First, the branches having the minimum utility are found (not selected) i.e.:

$$X_g^1 = \min\{U_g^{q1}(U_{g-1}^{q1}), U_g^{q2}(U_{g-1}^{q1}), U_g^{q3}(U_{g-1}^{q1})\}, \tag{5.1}$$

$$X_g^2 = \min\{U_g^{q1}(U_{g-1}^{q2}), U_g^{q2}(U_{g-1}^{q2}), U_g^{q3}(U_{g-1}^{q2})\}, \tag{5.2}$$

$$X_g^3 = \min\{U_g^{q1}(U_{g-1}^{q3}), U_g^{q2}(U_{g-1}^{q3}), U_g^{q3}(U_{g-1}^{q3})\}. \tag{5.3}$$

Next, the $(g-1)^{th}$ RN is going to select its move knowing $X_g^1$, $X_g^2$ and $X_g^3$. In SPE, the most dominating strategy for the $(g-1)^{th}$ RN is to choose a branch that improves the total utility of this sub-game i.e:

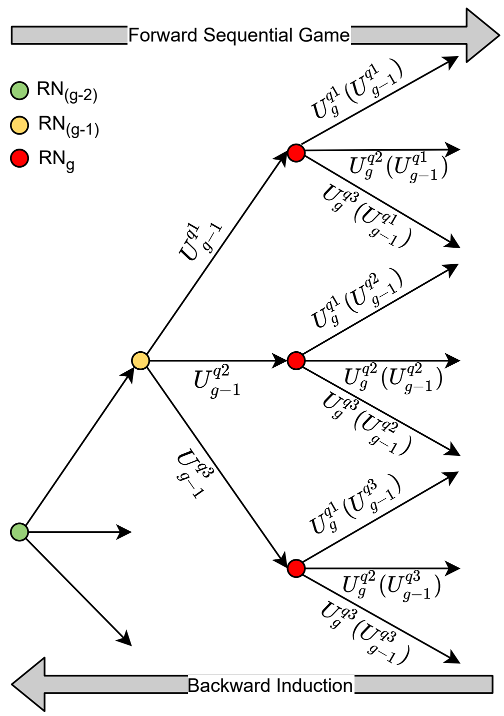



Figure 5.1: Sub-game perfect equilibrium between $g^{th}$ and $(g-1)^{th}$ RN during backward induction.

$$X_{g-1} = \min\{(U_{g-1}^{q1} + X_g^1), (U_{g-1}^{q2} + X_g^2), (U_{g-1}^{q3} + X_g^3\}), \tag{5.4}$$

where, $X_{g-1}$ is the SPE of the sub-game played between the $(g-1)^{th}$ and $g^{th}$ RNs. Similarly, this $X_{g-1}$ can be used by the $(g-2)^{th}$ RN in order to find its SPE, $X_{g-2}$, for the sub-game played between the $(g-2)^{th}$ and $(g-1)^{th}$ RNs. The same procedure is followed by each RN until the backward induction reaches the first RN that is $g = 1$. In Section 5.6, examples of backward induction are presented in the form of simulations.

## 5.5 Proof: A sequential game converges to Nash Equilibrium with SPE using backward induction

In game theory, Nash Equilibrium is a state of stability where there is no better strategy left for any player to play. In a sequential game, which has to be played in the forward direction (where, the players, do not have all the information of future/next players available), the players cannot reach a state of Nash Equilibrium during the forward iteration.

The sequential game can then be analyzed and converged to its Nash Equilibrium by finding Nash equilibria of its sub-games (each having at least two players) using backward induction i.e., iterating backwards in the game. This time, the prior players are enlightened with future information that was previously unavailable in the forward iteration.

In other words, with reference to the illustrative section above (5.4), $X_{g-1}$ is the Nash Equilibrium of the sub-game played between $(g-1)^{th}$ and $g^{th}$ RN. Moving backward, $X_{g-1}$ is used in finding the Nash equilibria of the sub-game played between the $(g-2)^{th}$ and $(g-1)^{th}$ RNs, $X_{g-2}$, eventually converging towards the Nash Equilibrium of the overall game. In conclusion, it can be said that Nash Equilibrium of a sequential game is equal to the overall Nash of its sub-games.

Table 5.1: Example 1 and Example 2 simulation parameters with random channel environments.

| **Parameters** | **Values (Example1)** | **Values (Examples2)** |
|---|---|---|
| $D_s r1$ | 172 Mbps | 91.2 Mbps |
| $f_s r1$ | 4.87 GHz | 4.87 GHz |
| (P,f) | 0.001, 0.25, 0.5 (W),<br>1.2 G, 850 M, 1.68 G (Hz) | 0.001, 0.25, 0.5 (W),<br>860 M, 2.75 G, 5.25 G (Hz) |
| $P_{g,avail}$ | 0.3, 0.2, 0.05 (W) | 0.2, 0.3, 0.1 (W) |
| $\mathrm{buffsize_g}$ | 2383, 2042, 2251 (bits) | 2476, 2315, 2239 (bits) |
| $d_g$ | 477, 359, 165 (m) | 413, 211, 193 (m) |

## 5.6 Performance Evaluation

In this section, the performance of proposed game theoretic approach is evaluated, using computer-based simulations. It is also compared with the centralized oracle, distributed greedy heuristic, CA, PACA, and random channel allocation algorithms. MATLAB was used to construct simulations using the following parameters: $F = 3000$ (bits), $SNR_{min} = -5$dB and $BW = 20$MHz.

With reference to Sections 5.4 and 5.5, Tables 5.2-5.4 demonstrate simulation results for two sample scenarios where Q=9 and G=3. The simulation parameters for these examples are shown in Table 5.1, where, $g = 1, 2, 3$. Tables 5.2 and 5.5 list the values for $po$ for $RN_1$ and the corresponding $po$ of $RN_2$ on the active ($po = 1$) $RN_1$ branches. Moreover, these tables also represent the corresponding $X_3$ for $RN_3$ and $U_2$ for $RN_2$.

Table 5.2: Backward Induction Example 1.

| **po** ($RN_1$) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| **po** ($RN_2$) | | | | | | | | |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| $U_2$ | | | | | | | | |
| 0 | 0 | 0 | 7.99e-05 | 0 | 0 | 9.15e-05 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1.59e-05 | 0 | 0 | 1.65e-05 | 0 |
| 9.39e-05 | 0 | 0 | 0 | 0 | 0 | 9.15e-05 | 0 | 0 |
| 0 | 1.66e-05 | 0 | 0 | 0 | 0 | 0 | 1.65e-05 | 0 |
| 0 | 1.66e-05 | 0 | 0 | 1.59e-05 | 0 | 0 | 0 | 0 |
| $X_3$ | | | | | | | | |
| 0 | 0 | 0 | 6.79e-05 | 0 | 0 | 6.58e-05 | 0 | 0 |
| 0 | 0 | 0 | 0 | 6.58e-05 | 0 | 0 | 6.58e-05 | 0 |
| 6.58e-05 | 0 | 0 | 0 | 0 | 0 | 6.58e-05 | 0 | 0 |
| 0 | 6.58e-05 | 0 | 0 | 0 | 0 | 0 | 6.58e-05 | 0 |
| 0 | 6.58e-05 | 0 | 0 | 6.58e-05 | 0 | 0 | 0 | 0 |

Table 5.3: Example 1: Backward Induction converging to better results than prior forward iteration.

| | $X_1$ | $U_2$ | $X_2$ | $X_3$ |
|---|---|---|---|---|
| | 2.14e-04 | 6.60e-05 | 7.99e-05 | 6.79e-05 |
| **Backward Induction** | 9.69e-05 | 1.51e-05 | 1.59e-05 | 6.58e-05 |
| | 2.19e-04 | 6.24e-05 | 9.15e-05 | 6.58e-05 |
| **Forward Game** | 9.73e-05 | 1.48e-05 | 1.65e-05 | 6.58e-05 |
| | 9.81e-05 | 1.62e-05 | 1.59e-05 | 6.58e-05 |

Tables 5.3 and 5.4 exhibit the results achieved using backward induction, where $X_2$ is said to be the Nash Equilibrium of the sub-game between $RN_2$ and $RN_3$ and $X_1$ is the Nash of the Nash for the sub-game between $RN_1$ and $RN_2$. Example 1 (Table 5.3) shows results for the case where backward induction converges to choose better branches as compared with the forward game iteration. Example 2 (Table 5.4) illustrates the case where backward induction results in choosing the same branches as the ones picked during the forward game iteration.

Table 5.5: Backward Induction Example 2.

| **po** ($RN_1$) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| **po** ($RN_2$) | | | | | | | | |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| $U_2$ | | | | | | | | |
| 0 | 0 | 0 | 6.45e-05 | 0 | 0 | 1.41e-04 | 0 | 0 |
| 0 | 0 | 0 | 6.45e-05 | 0 | 0 | 0 | 0 | 0 |
| 3.92e-05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.92e-05 | 0 | 0 | 6.45e-05 | 0 | 0 | 0 | 0 | 0 |
| 3.92e-05 | 0 | 0 | 6.45e-05 | 0 | 0 | 0 | 0 | 0 |
| $X_3$ | | | | | | | | |
| 0 | 0 | 0 | 9.03e-05 | 0 | 0 | 2.67e-04 | 0 | 0 |
| 0 | 0 | 0 | 9.03e-05 | 0 | 0 | 0 | 0 | 0 |
| 2.67e-04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2.67e-04 | 0 | 0 | 9.03e-05 | 0 | 0 | 0 | 0 | 0 |
| 2.67e-04 | 0 | 0 | 9.03e-05 | 0 | 0 | 0 | 0 | 0 |

Table 5.4: Example 2: Forward game and backward induction, both returning the same branches.

| | $X_1$ | $U_2$ | $X_2$ | $X_3$ |
|---|---|---|---|---|
| | 2.45e-04 | 9.04e-05 | 6.45e-05 | 9.02e-05 |
| **Chosen Branch** | 1.71e-04 | 1.652e-05 | 6.45e-05 | 9.02e-05 |
| | 3.26e-04 | 1.99e-05 | 3.92e-05 | 2.67e-04 |
| | 2.96e-04 | 1.41e-04 | 6.45e-05 | 9.02e-05 |
| | 1.73e-04 | 1.85e-05 | 6.45e-05 | 9.02e-05 |

In Fig. 5.2, the proposed greedy, oracle, and game-theroetic approaches are compared with random channel assignment and the conventional CA and PACA approaches for $Q = 100$, $50 < d_g(m) < 1000$, $0.05 < P_{avail}(W) < 0.5$, $100 < \mathrm{buffsize_g(bits)} > 300$, $0.001 < P(W) < 0.5$.

It can be seen from the referred figure that in random channel allocation, the RNs select channels randomly (without any strategy), leading them to take the longest time to relay the data packet. Furthermore, it can be noticed that PACA algorithm performs better than many other approaches but still has poor performance as compared with the proposed game model. The CA algorithm performs worse than most approaches; however, it is still superior in performance to the random and oracle approaches.

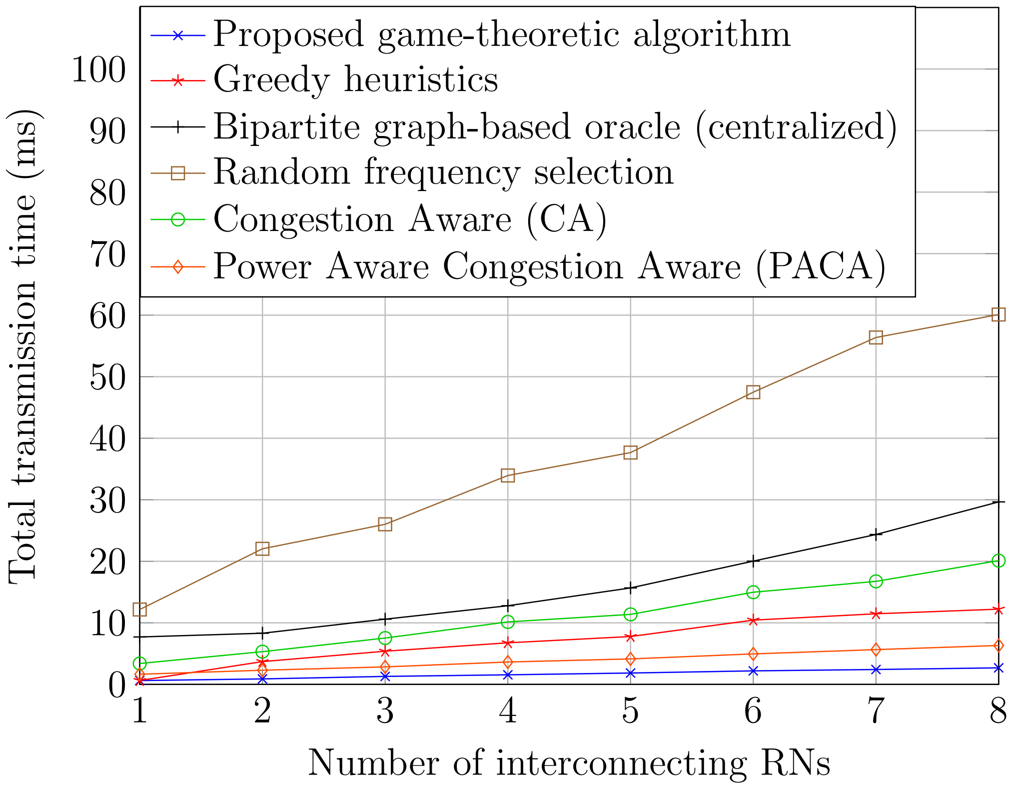



Figure 5.2: Comparison of the total transmission time for the considered methods with a varying number of RNs.

In Fig. 5.3 the buffer overflow results are provided, for greedy, game, and oracle approaches as compared with the random, CA and PACA ones, for $G = 5$, $Q = 20$, $50 < d_g(m) < 100$, $1.5 < P_{avail}(W) < 3$, $\text{buffsize}_\text{g}(\text{bits}) = 100, 200, 300$, $0.01 < P(W) < 3$.. It can be seen from the figure that even in the worst-case scenario, the proposed game model still provides a minimum number of buffer overflow bits as compared with all other approaches. It can also be noticed that the oracle approach has a smaller buffer overflow than all other approaches save for that of the proposed game model. However, the PACA and CA algorithms lie in the middle for this criteria.

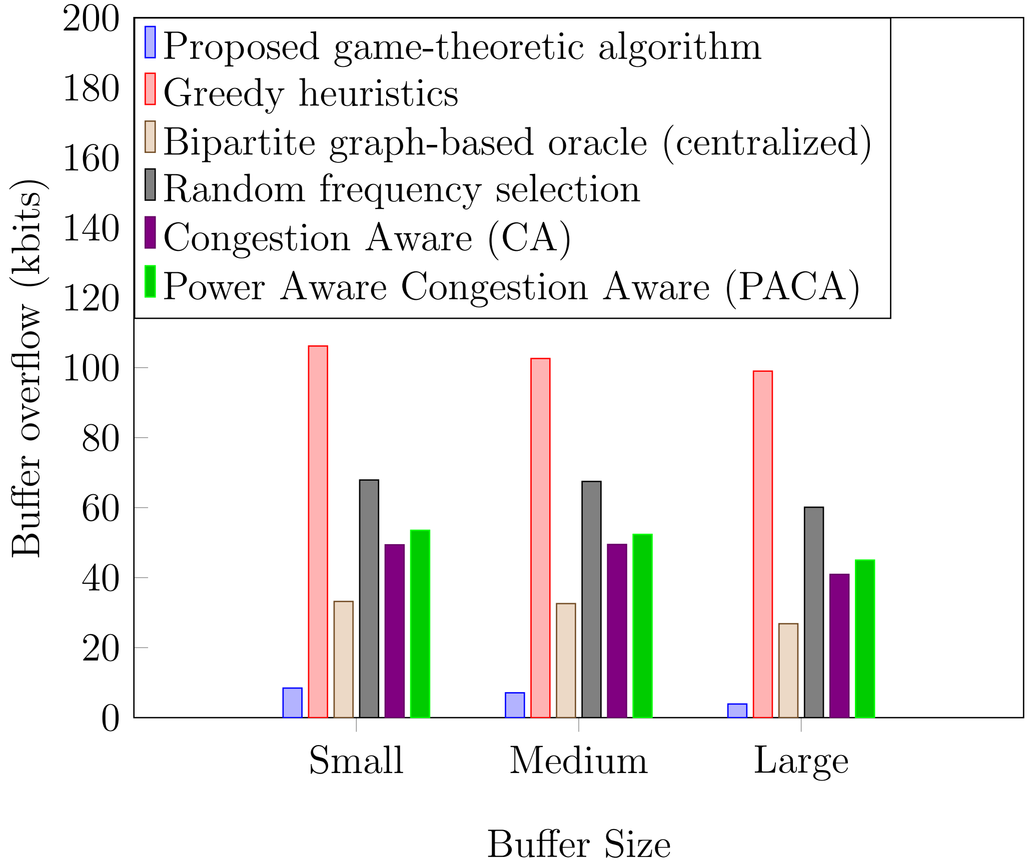



Figure 5.3: Buffer overflow comparison with different buffer sizes for five RNs.

In Fig. 5.4 the effective throughput of the approaches is provided for $G = 4$, $Q = 100$, $50 < d_g(m) < 1000$, $0.05 < P_{avail}(W) < 0.5$, $100 < \mathrm{buffsize_g(bits)} > 3000$, $0.001 < P(W) < 0.5$. It can be seen from the figure that even in the worst-case scenario, the proposed game model has the highest effective throughput, whereas, the random channel allocation has the least effective one because of its random strategy selection. The CA algorithm performs better than oracle and random ones but still has a smaller effective throughput than that of the game, greedy, and PACA models. The PACA algorithm per-

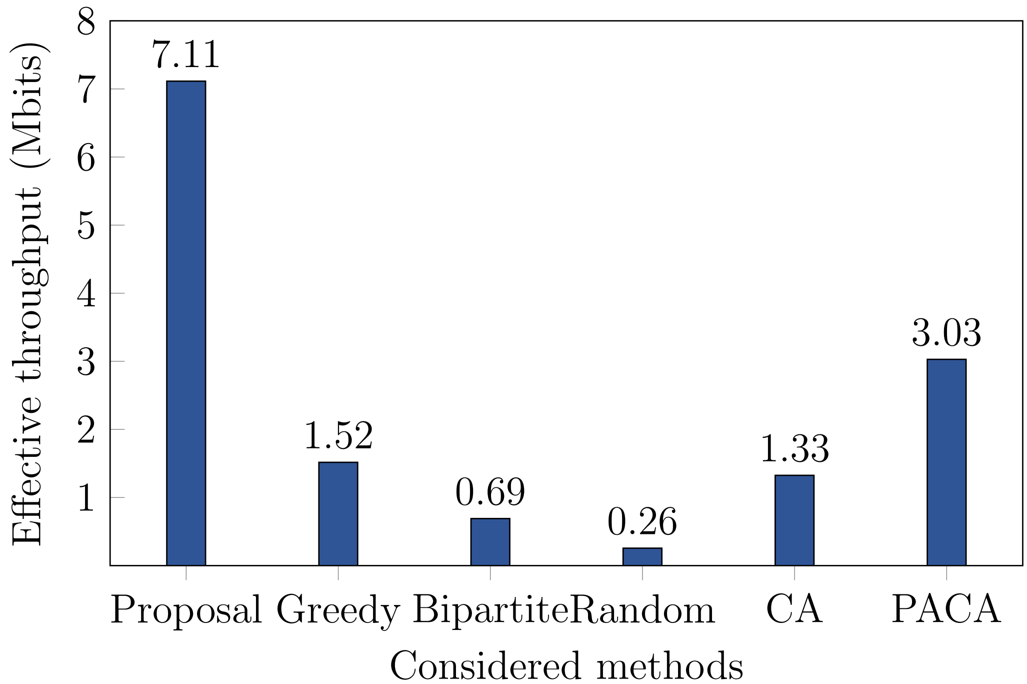



Figure 5.4: Effective throughput comparison with different buffer sizes for 4 RNs.

forms better than other approaches but still has smaller effective throughput as compared with the proposed game model.

## 5.7 Conclusion

In this chapter, model-specific Nash Equilibrium analysis is provided for the proposed game model. The results from simulations demonstrate that the performance of proposed model eclipses that of traditionally used approaches, including a centralized oracle and a greedy heuristic-based distributed method.

## 5.8 Publications Resulted from This Chapter

- B. Mughal, Z. M. Fadlullah, M. M. Fouda and S. Ikki, "Applying Game Theory to Relay Resource Selection in Hybrid-band Wireless Systems," Submitted to *IEEE Sensors Journal*, Nov. 2021.

# Chapter 6

# Optimizing Packet Forwarding Performance in Multi-Band Relay Networks via Customized Reinforcement Learning

**Motivation:** The centralized decision algorithms are typically computationally hard and require non-deterministic polynomial time. The use of such approaches in literature, fail to consider the rapidly changing channel conditions. This motivates us to explore a distributed self learning algorithm, known to be ML techniques, in order to propose smart and adaptive decision approach.

## 6.1 Introduction

Since multi-band channel allocation to RNs is a time-sensitive problem, a suitable (preferably optimal) decision is required without introducing more delay to the communication while solving the problem. As depicted in Fig. 6.1, the frequency channels are to be decided for RNs to forward the data to the DN. If considered centrally, there needs to be one centralized oracle which takes channel allocation decisions for the entire network on the basis of prior channel environment conditions received from RNs.

In contrast, in a distributed approach, the RNs need to take individual decisions in accordance with their local channel environment. The centralized approach requires complete information and coordination from/with RNs to provide a one-shot (static) optimal solution, which are more likely invalid by the time the central oracle decides them. Such an approach is impractical and not suitable for a time-sensitive dynamically changing networks. Hence, the preference hinges towards using distributed decision approach for solving the problem of multi-band channel allocation to RNs, providing reliable, fast, and valid decisions.

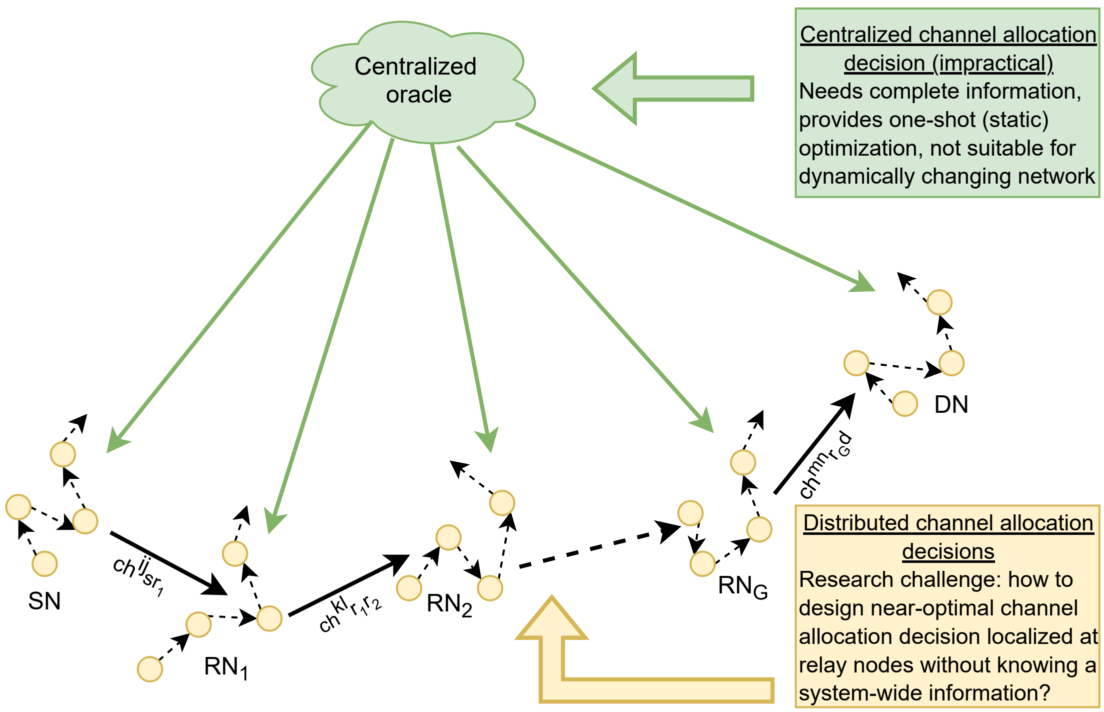



Figure 6.1: Centralized vs distributed channel allocation in mobile multi-band relay network.

The previous work considered distributed algorithms [33, 39] to solve this problem with the assumption of stable channel conditions and static radios. In this chapter, a dynamic relay network scenario is considered, with unstable channel environments and mobile radios. This requires a new distributed solution to account for the complex, highly dynamic relay network environment.

The original, computationally hard problem is reformulated as a MDP [40]. Since solving the MDP to derive optimal multi-band channel allocation delay by individual relay nodes is still an expensive process, due which its corresponding reinforcement learning problem is designed, which is then solved using a customized Q-Learning algorithm with epsilon-greedy policy [7].

***Contributions:*** With the objective of minimizing the packet latency in multi-band relay network, the contributions of the work, in this chapter, are as follows.

1. An optimization problem for multi-band channel allocation under mobile and dynamic radio environment is proposed, and its computationally hard nature and need for complete information is demonstrated. Hence, the original problem is translated as an MDP, and provide details on how the proposed model satisfies MDP properties.

2. The proposed MDP is redesigned as a reinforcement learning-based problem where details on designed states, actions and rewards are provided. A customized Q-learning algorithm is proposed so as to solve the reformulated problem.

3. Convergence bounds are provided for proposed algorithm through empirical results by fine-tuning the relevant learning parameters, and also empirical performance comparison of proposed method with comparable, conventional methods.

The remainder of the chapter is structured as follows. In section 6.2, the considered system model is presented along with an optimization problem formulation. In section 6.3, preliminaries on MDP are provided where, the original optimization problem is reformulated into an MDP. Section 6.4 presents the proposed Q-Learning-based reinforcement learning algorithm. The performance of the proposal is evaluated and compared with conventional methods in section 6.5. Finally, the chapter is concluded in section 6.6.

## 6.2 System Model and Problem Formulation

In this section, considered system model is provided for multi-band relay-based network, which leads to discussing optimization problem for assigning the channels of different bands to RNs. For ease of reference for the readers, the major notations and symbols used throughout this section and the remainder of the chapter are listed in Table 6.1. The assumptions are presented as follows:

**First**, some mobile terminals have already been selected as RNs, and the incentive/policy used by the network operator allowing some mobile users to act as RNs is beyond the scope of the current work.

**Second**, spectrum sensing [11, 41] is considered to be performed by the radios before accessing channels.

**Third**, the SNR estimation [30] is assumed to be carried out by the nodes before taking decisions.

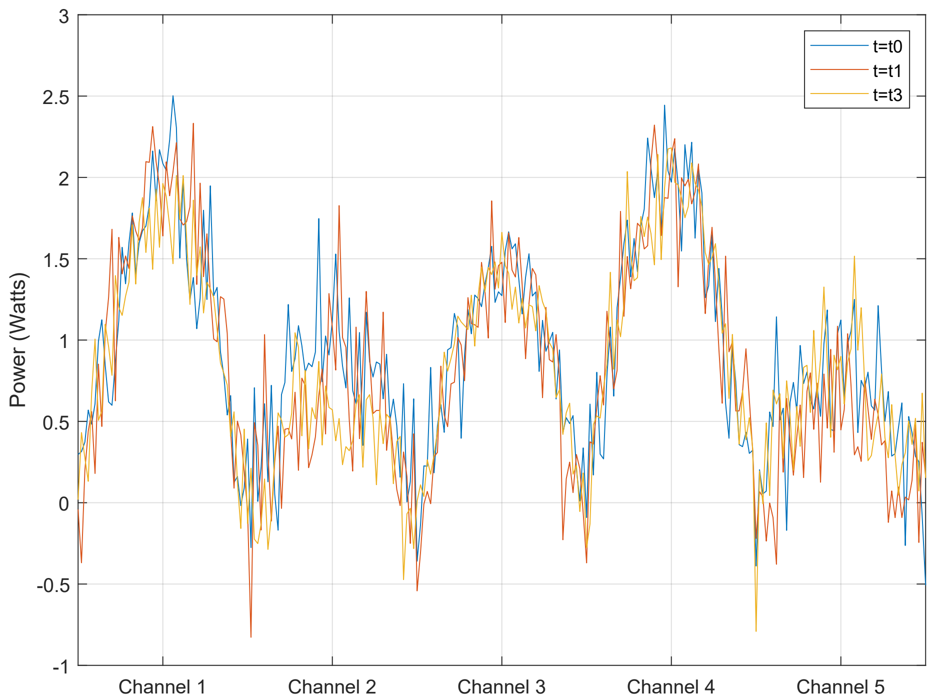
Any RN selection approach [42], spectrum sensing algorithm, and SNR estimation techniques can be used for the practical implementation of proposed algorithm. Note that the joint channel equalization and estimation for multiple frequency bands is still an open research issue.

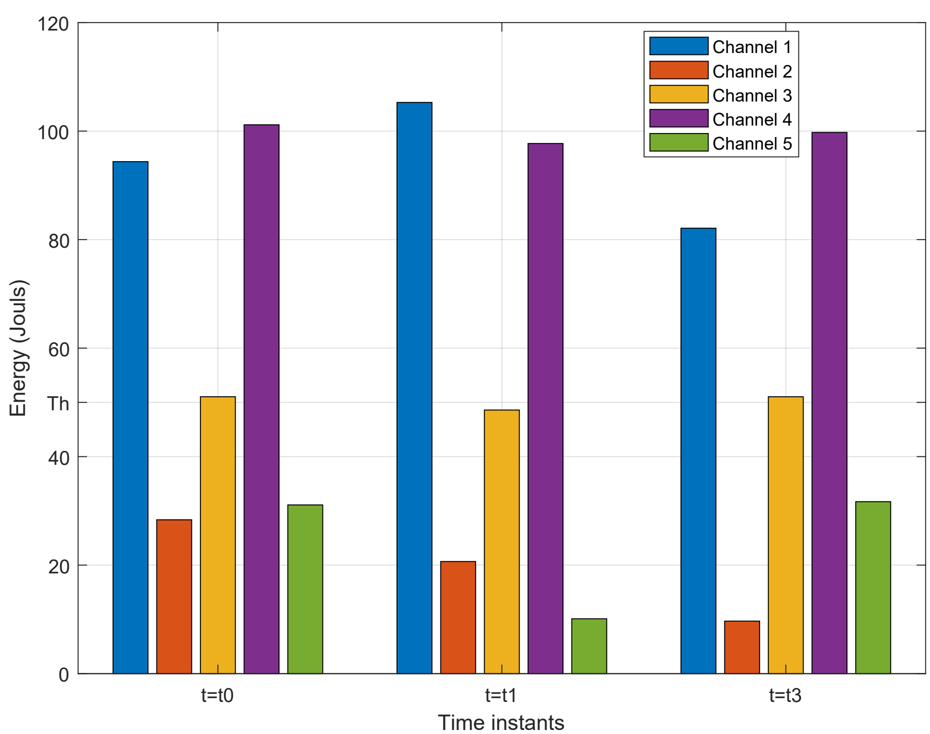
Since this is also beyond the scope of this research, it is to be considered that the channel estimation information is available at the RNs, because this is assumed to have a constant and relatively negligible effect on the packet latency in the considered relay network.

With the above assumptions, the considered system model is now described, with an example of the channel dynamism under AWGN as depicted in Fig. 6.2a. Its corresponding energy detection results are demonstrated in Fig. 6.2b.

Table 6.1: List of major technical notations and symbols used in the paper.

| Notations | Description |
|---|---|
| $F_{sr}$ | Data frame sent from SN to RN |
| $F_{rd}$ | Data frame from RN to DN |
| $t_{ir}$ | Time at which RN starts receiving $F_{sr}$ |
| $t_{out}$ | Time at which RN schedules the data frame for re-transmitting it to DN ($F_{rd}$). |
| $T_h$ | Time span of the header of the packet. |
| $T_{sr}$ | Time span of $F_{sr}$. |
| $T_{rd}$ | Time span of $F_{rd}$. |
| $T_{tot}$ | Total time span from the instant at which RN starts receiving data frame from SN ($F_{sr}$) till the time instant, DN completely receives the re-transmitted data frame ($F_{rd}$) from RN. |
| $\mathcal{D}_{sr}$ | Rate at which SN transmits data frame to RN ($F_{sr}$). |
| $\mathcal{D}_{rd}$ | Rate at which RN re-transmits data frame to DN ($F_{rd}$). |
| $ch_{sr}^{ij}$ | i-th frequency channel of j-th frequency band used by SN to transmit $F_{sr}$ to RN. |
| $ch_{rd}^{mn}$ | m-th frequency channel of n-th frequency band used by RN to re-transmit $F_{rd}$ to DN. |
| $\mathrm{SNR_{sr}}$ | Estimated SNR of the channel used by SN to transmit $F_{sr}$ to RN. |
| $\mathrm{SNR_{rd}}$ | Estimated SNR of the channel used by RN to re-transmit $F_{rd}$ to DN. |
| buffsize | Buffer size available to the node. |
| $f_{rd}$ | Frequency channel used by the RN to forward data packet to its next destined node. |
| $f_{sr}$ | Frequency channel used by the source node to transmit data packet to the RN. |
| $\alpha$ | Learning rate of Q-Learning algorithm. |
| $\gamma$ | Future reward parameter for future state/action. |
| $\epsilon$ | Parameter for probability of choosing a random action. $(1-\epsilon)$, is probability of choosing greedy (best) action). |




(b) Energy detection on channels where Th. stands for threshold.

Figure 6.2: Spectrum sensing example.

Based on this considered relay network dynamism, the preferred distributed decision approach is now discussed. Subscripts $sr$ and $rd$ are used for any pair of SN, RN and DN. For any particular node, the previous node becomes source node and the next node becomes the DN. Once any RN start receiving a data packet $F_{sr}$, at time $t_{ir}$, at channel $f_{sr}$, with data rate $\mathcal{D}_{sr}$, from its prior/source node (can be one of the RNs or the SN itself), the RN, being able to communicate on multi-bands, can now commence forwarding the data packet $F_{rd}$ at $t_{out}$, at data rate $\mathcal{D}_{rd}$, to its next destined node using another channel $f_{rd}$ from a different band.

At this point, RN needs to decide the feasible $\mathcal{D}_{rd}$, $t_{out}$ and $f_{rd}$ values suitable for the incoming data rate. The selection of $\mathcal{D}_{rd}$, $t_{out}$ and $f_{rd}$ depends on the resources available at RN that include available power $P_{avail}$, size of the buffer (buffsize), and availability/occupancy of frequency channels. In case where, the RN finds that all channels are occupied, it forwards packets using the DF scheme (that is on the same channel, $f_{rd} = f_{sr}$) by finding the best possible $\mathcal{D}_{rd}$. In this case:

$$T_{tot} = T_{sr} + T_{rd}, \tag{6.1a}$$

$$t_{out} = t_{ir} + T_{sr}, \tag{6.1b}$$

$$BO = F_{sr} - \text{buffsize}, \tag{6.1c}$$

where $T = F(\mathcal{D})^{-1}$, $T_{tot}$ denotes the total time needed for RN to complete the simultaneous reception and re-transmission of $F_{sr}$ and $F_{rd}$, respectively. BO represents the observed buffer overflow in cases where buffsize is limited.

In the case where RN finds the available channels with better SNR than SNR of receiving channel ($\mathrm{SNR_{rd}} > \mathrm{SNR_{sr}}$), the data packets can now be forwarded at a higher data rate ($\mathcal{D}_{rd} > \mathcal{D}_{sr}$). This results in $T_{rd} < T_{sr}$. In this case:

$$T_{tot} = T_{sr}, \tag{6.2a}$$

$$t_{out} = t_{ir} + T_{sr} - T_{rd}, \tag{6.2b}$$

$$BO = [(T_{rd} - (T_{sr} - T_h))\mathcal{D}_{rd}] - \text{buffsize}, \tag{6.2c}$$

where $T_h$ is the header time of $F_{sr}$. When the channel has lesser $\mathrm{SNR_{rd}}$ than $\mathrm{SNR_{sr}}$, the packet can be forwarded at a lesser $\mathcal{D}_{rd}$ than $\mathcal{D}_{sr}$. This results in $T_{rd} > T_{sr}$. Here,

$$T_{tot} = T_{rd} + T_h, \tag{6.3a}$$

$$t_{out} = t_{ir} + T_h, \tag{6.3b}$$

$$BO = [(T_{sr} - T_{rd})\mathcal{D}_{sr}] - \text{buffsize}. \tag{6.3c}$$

With the aforementioned system model, the multi-band channel assignment can now be reformulated as a minimization problem with the following objective function (6.4a) under several constraints (6.4b-6.4d). Note that constraint 6.4c signifies the highly dynamic relay conditions since the SNR of the links between the relay nodes (or relay-destination nodes) is treated as a function of the link power, frequency bands dynamics, and the changing distance under mobility effect.

$$\min_{\mathcal{D}_{rd}, f_{rd}} (t_{out} + T_{rd}), \qquad s.t. \tag{6.4a}$$

$$BO(\mathcal{D}_{sr}, \text{buffsize}) \leq \text{BO}_{\text{min}}, \tag{6.4b}$$

$$\text{SNR}_{\text{rd}}(\text{P}_{\text{rd}}, \text{f}_{\text{rd}}, \text{dist}_{\text{rd}}) \geq \text{SNR}_{\text{min}}, \tag{6.4c}$$

$$P_{rd} \leq P_{avail}, \tag{6.4d}$$

where $BO_{min}$ is minimum allowed value of overflown buffer, $\text{SNR}_{\text{min}}$ is the QoS constraint, $P_{rd}$ is the power required to transmit at $\mathcal{D}_{rd}$ and $dist_{rd}$ is the distance between the RN and its destined next node.

Based on the earlier work in [39], this optimization problem can be treated as a computationally hard (NP) problem that cannot be solved in polynomial time for a relay network with a large number of RNs, frequency bands, and channels.

Furthermore, the complete information across the entire relay topology is required by a central oracle to solve this problem even for a relatively small search space. Due to this practicality issues, it is to be explored that how to remodel this optimization problem in a distributed manner in the following section.

## 6.3 Problem Reformulation with MDP

In this section, the aim is to reformulate the original optimization problem as a MDP from the perspective of a RN. This is required for a distributed optimal decision regarding the multi-band channel allocation at the RN level given only the local information available to the RN. First the preliminaries on MDP are provided, and then the problem reformulation as an MDP is discussed.

### 6.3.1 MDP Preliminaries

The main characteristic which defines a Markov process is that the future states depend on the current state only. For example, if at a certain time, an agent is in state $S_t$, a random action is being taken $A_t$, now the occurrence of the next state $S_{t+1}$ depends on this current action, $A_t$, only and not on action taken at the previous instant, $A_{t-1}$.

Above is the property of any process to be defined as a Markov process. the proposed problem of RN choosing best decision of forwarding data packet is modeled as an MDP and is sculptured as to learn from the random actions taken in certain states.

For any system, these actions and states have certain values, which are defined by Bellman's equations and can also be observed from backup diagrams as shown in Fig. 6.3. Following is the Bellman equation for value of a state following policy $\pi$, denoted as $V_\pi$:

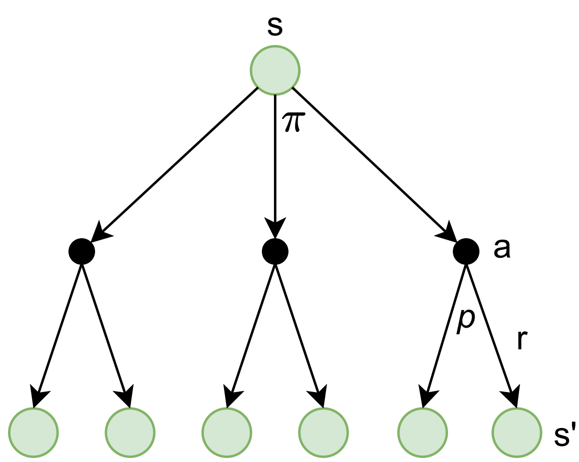
$$V_\pi(S) = \sum_A \pi(A|S) \sum_{S'} p(S'|A)[R + \gamma V_\pi(S^{'})], \tag{6.5}$$

where $\alpha$ represents the learning parameter, which impacts how fast the algorithm must learn while $\gamma$ indicates the future reward parameter to give importance to the value of the next state, $V_\pi(S^{'})$.

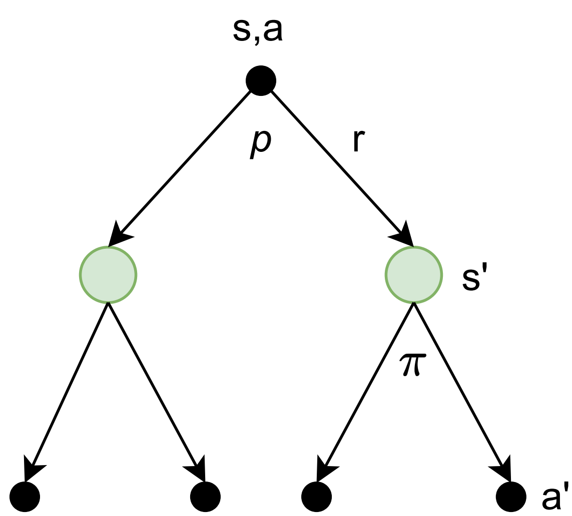
Through this equation, the values of a current and subsequent state(s) can be related as depicted in Fig. 6.3a. This can be regarded as foreseeing the possible states in the future with respect to the current state, each of which is denoted by an open circle. The state-action pairs are represented by solid circles.

As shown in Fig. 6.3a, the top (root) node signifies state $s$, from which any of the three set of actions could be taken by adhering to a policy $\pi$. The environmental response could be a further sequence of succeeding steps that can be measured by a reward.

The reward calculation is done a function $p$. By averaging over all possible states and weighing each state in terms of its probability, eq. (6.5) ascertains that the beginning state value equals the sum of the expected next state and the expected reward.




(a) Backup diagram for value of a state while following policy $\pi$.




(b) Backup diagram for value of action when following policy $\pi$.

Figure 6.3: Backup diagrams demonstrating the MDP for values of a state and an action, respectively, subject to a given policy.

Similarly, following is the Bellman equation for value of an action $A$ taken according to policy $\pi$ from the state $S$, donated as $Q_\pi(S, A)$:

$$Q_\pi(S, A) = \sum_{S'} p(S^{'}|A)[r + \gamma V_\pi(S^{'})] \tag{6.6}$$

Imagine the solid circle on the top in Fig. 6.3b, as an action taken from some state $S$. This action taken now has different probabilities $p$ of resulting in either of the next two states shown in the figure. The Bellman equation for an action, averages over all the possibilities of its successor states and rewards, prioritizing each according to its likelihood. Thus, eq. (6.6) represents the beginning state value as the sum of the anticipated next state and the associated reward.

### 6.3.2 Original Problem Reformulation into MDP

Based on the preliminaries, the original problem (eq. (4)) can now be formally transformed into a distributed, finite-state MDP as depicted in Fig. 6.4. the designed states, actions, and rewards are demonstrated in the figure. The MDP model comprises seven states, denoted by $\boldsymbol{S} = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7\}$, where, a RN is an agent taking actions and interacting with the environment under dynamically changing channel conditions of multiple frequency bands and nodes mobility. The set of states, and their corresponding sets of next states, actions and rewards are discussed in detail as follows.

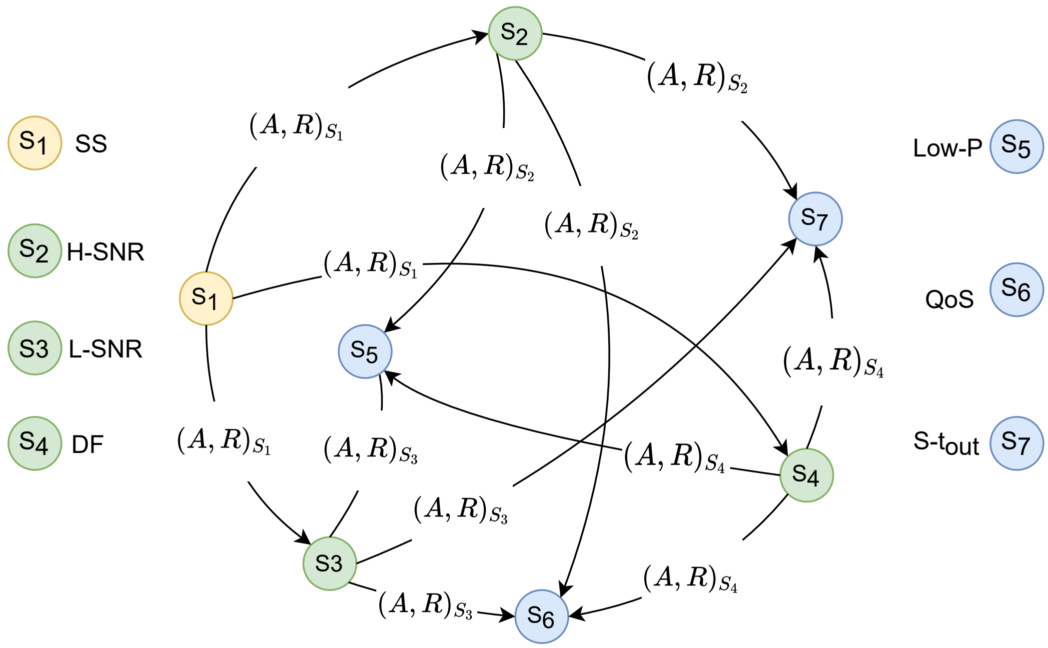



Figure 6.4: State diagram for proposed finite state MDP reinforcement learning algorithm.

1. $S_1$ is the **initial state** of RN where, RN performs **spectrum sensing** along with **SNR estimation** on a certain frequency channel from any band.

   Hence, the **action set** of $S_1$ is to choose the frequency channel $f$ to be sensed and estimated given as: $\boldsymbol{A_{s_1}} = \{f_1, f_2, ..., f_{\mathcal{F}}\}$, where $\mathcal{F}$ delineates the total possible actions available for RN in this state, which is the set of channels that RN, being a secondary user, is allowed to use only after sensing. In this chapter, $S_1$ is to be referred as the **SS-state**.

   The associated **set of rewards** for this state is given as $\boldsymbol{R_{s_1}} = \{R_1^{s_1}, R_2^{s_1}, R_3^{s_1}\}$, such that $R_1^{s_1} > R_2^{s_1} > R_3^{s_1}$. Once RN takes an action of picking say, $f_1$, and senses it as free with high SNR, the environment needs to assign $R_1^{s_1}$ against action $f_1$ leading

RN to the next state of $S_2$. $R_2^{s_1}$ is to be assigned to the action $f_1$ with next state as $S_3$, in case where RN finds this channel to be free, but with lesser SNR. $R_3^{s_1}$ is to be associated with the action if the sensed channel is occupied, leading the agent to the next state of $S_4$. The set of next states for this state is given as $\boldsymbol{S_1'} = \{S_2, S_3, S_4\}$.

2. $S_2$ is an **intermediate state** of the proposed MDP model, where RN has reached from the previous state of $S_1$ taking an action of picking a frequency channel which is free and has higher SNR. This state is referred to as the **H-SNR-state**. In this state, the RN takes an action of choosing data rate from a standard set of data rates, which are achievable on this high SNR free channel.

   The **action set** for this state can be given as $\boldsymbol{A_{s_2}} = \{\mathcal{D}_1^{s_2}, \mathcal{D}_2^{s_2}, ..., \mathcal{D}_{\mathcal{H}}^{s_2}\}$, where $\mathcal{H}$ are the total number of data rates achievable at this state. As can be seen from Figs. 6.5 and 6.6, the data rates chosen in this state will not have impact on the total time of simultaneous reception and transmission, hence more value is given to the data rate using the least power.

   Hence, the **set of rewards** for this state is the value inversely proportional to the data rate chosen that is $\boldsymbol{R_{s_2}} = \{R_1^{s_2} \propto (\mathcal{D}_1^{s_2})^{-1},\ R_2^{s_2} \propto (\mathcal{D}_2^{s_2})^{-1}, ....,\ R_{\mathcal{H}}^{s_2} \propto (\mathcal{D}_{\mathcal{H}}^{s_2})^{-1}\}$. If the selected data rate requires power more than the $P_{avail}$ of RN, the environment assigns zero or negative reward (also known as bad rewards) to the selected action and leads the agent to state $S_5$.

   On the other hand, if the selected data rate results in violating the QoS, the environment drives the agent to the next state of $S6$ by assigning bad reward to the action. However, if both constraints are met, the environment leads the agent to $S_7$ by assigning rewards according to the set $\boldsymbol{R_{s_2}}$. The set of next states for this state is given as $\boldsymbol{S_2'} = \{S_5, S_6, S_7\}$.

3. $S_3$ is an **intermediate state** of the proposed model where RN has reached from previous state of $S_1$ taking an action of picking a frequency channel which is free and has lesser SNR. In this chapter, this state is to be referred as **L-SNR-state**.

   In this state, the RN takes an action of choosing the data rate from a standard set of data rates, which are achievable on this low SNR free channel. The action set for this state can be given as $\boldsymbol{A}_{s_3} = \{\mathcal{D}_1^3, \mathcal{D}_2^3, ..., \mathcal{D}_{\mathcal{L}}^3\}$, where $\boldsymbol{\mathcal{L}}$ are the total number of data rates achievable at this state.

   As can be seen from Fig. 6.5, the data rates selected in this state will have a direct impact on total time of simultaneous reception transmission. Hence, more value is given to the data rate taking the least time. Thus, the **set of rewards** for this state is the value directly proportional to the data rate chosen, that is $\boldsymbol{R_{s_3}} = \{R_1^{s_3} \propto \mathcal{D}_1^{s_3},\ R_2^{s_3} \propto \mathcal{D}_2^{s_3}, ....,\ R_{\mathcal{L}}^{s_3} \propto \mathcal{D}_{\mathcal{L}}^{s_3}\}$.

   In this state, after an action is taken, the environment follows the same criteria for choosing the next states as discussed in case of $S_2$.

4. $S_4$ is an **intermediate state** of the proposed model, where RN has reached from the previous state of $S_1$ taking an action of picking a frequency channel which is not free. This state is to be called as **DF-state** in this chapter.

   Here, the **action set** is the union set of all the data rates $\boldsymbol{A_{s_4}} = \boldsymbol{A_{s_2}} \cup \boldsymbol{A_{s_3}}$, as seen from Fig. 6.6. In this state, after an action is adopted, the environment follows the same criteria for choosing the next states as discussed for $S_2$, and the rewards are assigned as directly proportional to the data rate (same as discussed in case of $S3$ as time is the significant element in this case).

5. $S_5$ is the **bad terminal state** of the proposed model, where RN reaches from any of $S_2, S_3, S_4$ by taking action (choosing data rate) requiring power more than $P_{avail}$. Here, this state is to be referred as **Low-P-state**. The actions causing this state are given poor rewards for assisting the further learning process.

6. $S_6$ is the **bad terminal state** of the proposed model, where RN reaches from any of $S_2, S_3, S_4$ by taking action (choosing data rate) resulting in SNR lesser than the minimum QoS requirement, $\text{SNR}_{\text{min}}$.

   This state is to be called as **QoS-state**. The actions causing this state are given poor rewards for helping the learning process.

7. $S_7$ is the **good terminal state** of the proposed model, where RN reaches from any of $S_2, S_3, S_4$ by taking action (choosing data rate) feasible in terms of $P_{avail}$ and $\text{SNR}_{\text{min}}$.

   The actions causing this state are given rewards according to their corresponding reward sets $R_{s_2}, R_{s_3}, R_{s_4}$. This state is to be referred as **S-$t_{out}$-state**, which means if RN reaches to this state, the packet can be scheduled at $t_{out}$.

While theoretically the proposed MDP-based reformulated problem model is elegant, estimating the transition probabilities between the states and resolving an optimal channel assignment is not trivial, and still is computationally expensive for an individual RN. Therefore, a distributed learning technique is needed to solve the MDP for optimal multi-band allocation to relay nodes, which is designed in the following section.

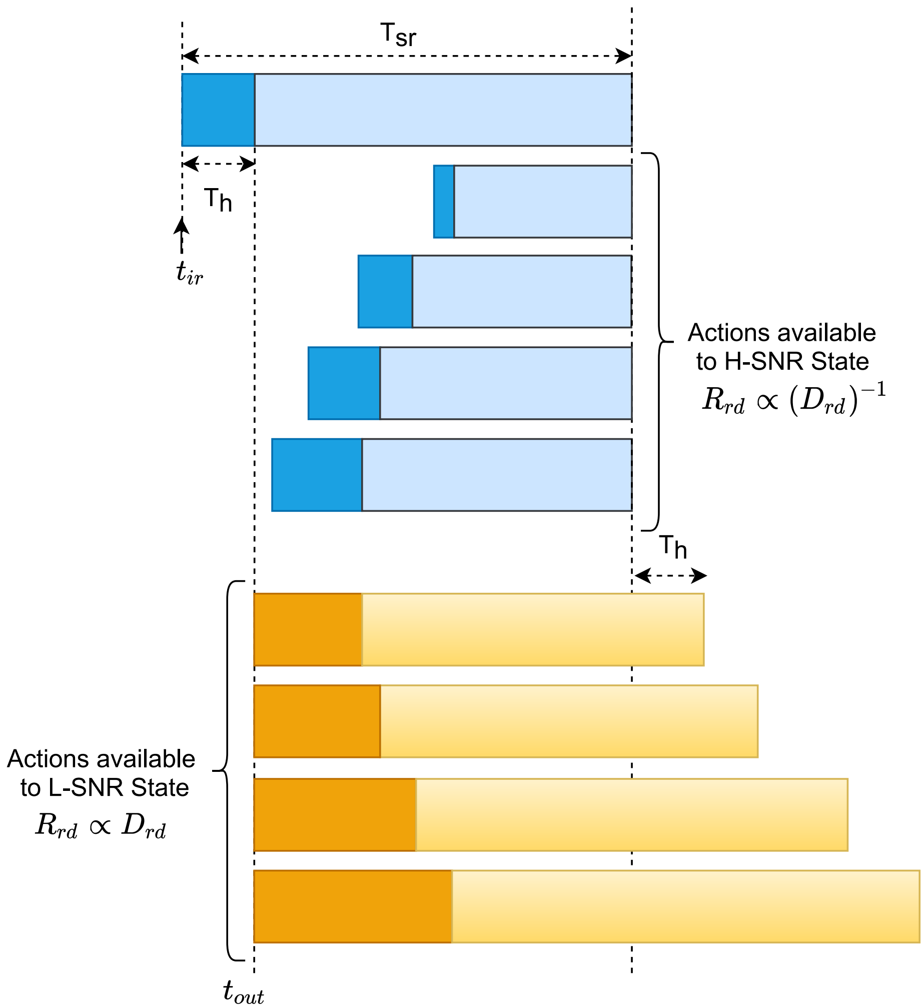



Figure 6.5: High-level view of action sets available to $S_2$ and $S_3$ and their corresponding rewards.

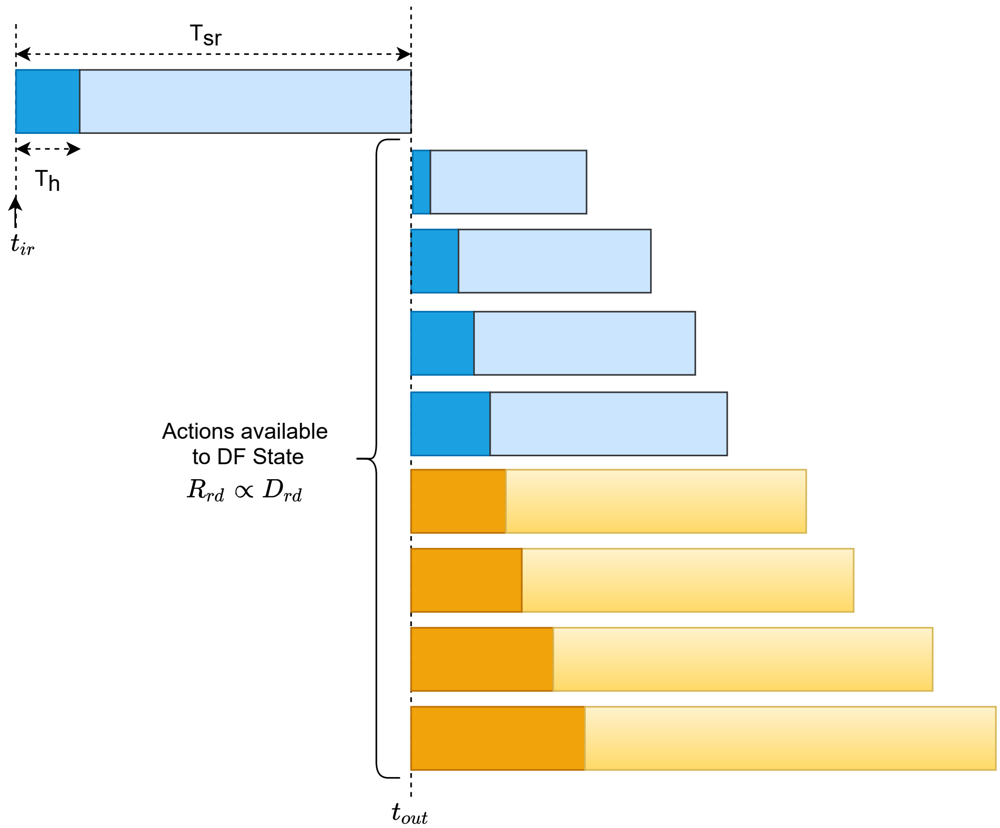



Figure 6.6: High-level view of action sets available to $S_4$ and its corresponding rewards.

## 6.4 Proposed Reinforcement Learning-Based Optimal Multi-Band Allocation to Relay Nodes

To solve the re-formulated MDP problem, in this section, the aim is to design a reinforcement learning algorithm for capturing the ongoing process of interaction between an agent (i.e., a RN) and its environment. The agent being in a certain state, $S$, takes an action, $A$, according to some policy; and it inquires with its environment as shown in Fig. 6.7. According to the action taken by the agent, the environment now assigns a particular reward, $R$, against that action and also decides the next state, $S'$ for the agent to move onto. With this continuous interactions, the RN acting as the agent keeps learning the best sates and actions for itself. The actions are taken on the basis of a certain policy that could be either totally random, or greedy or epsilon-greedy ($\epsilon$-greedy).

The transition from one state to the next has certain probabilities. These probabilities are known as transition probabilities ($p$) and can be written in the form of transition tables [7]. The reformulated problem is represented as a finite MDP reinforcement learning model, which consists of a stochastic 4-tuple of states, actions, rewards and transition probabilities $(S, A, R, p)$.

Next, the details on proposed distributed Q-Learning algorithm is presented, following $\epsilon$-greedy policy. In order to find the optimal value of previously discussed Bellman Eqs. (6.6) and (6.5) of any MDP, Q-Learning algorithm can be used for finding the optimal decision solutions.

In MDP first each and every possible action is taken so that to find the probabilities of their occurrence and only then by using Bellman equations, the final optimal value of any state and action can be found. Where as, Q-Learning updates the estimates, based on other (already) learned estimates, without waiting for a final outcome. Such phenomena is also known as bootstrapping. Eq. (6.7) is known to converge to the optimal value of Eq. (6.6) for any MDP reinforcement learning problem.

$$Q_(S, A) = Q(S, A) + \alpha[r + \gamma \underset{A}{max} Q(S', A) - Q(S, A)] \tag{6.7}$$

There are mainly three policies which could be followed for selecting an action, i.e., random, greedy, and $\epsilon$-greedy [7]. In random policy, the agent takes an action randomly at each state, thereby exploring the options. In the greedy policy, the agent always adopts the best action (with the highest Q(S,A) value), thus exploiting the best known option.

On the other hand, in $\epsilon$-greedy policy, the agent keeps balance between exploring and exploiting the options by selecting the best action with probability $(1-\epsilon)$ and take a random action with probability $\epsilon$. To ensure the best learning over time, proposed reinforcement learning solution is geared by employing the $\epsilon$-greedy policy to customized Q-learning, as follows:

Q-Learning algorithm is utilized as a continuous learning process with dynamic channel environment and mobile relay nodes. Referring to the steps in Algorithm 5 for continuous back-end Q-Learning algorithm, at the start of each episode, being a mobile node with varying channel conditions, RN updates its $dist_{rd}$, its spectrum sensing results and also calculates its updated SNR estimation results.

In one episode, RN starts taking actions from the initial state, $S_1$ and reaches its terminal state. Then another episode starts with changed/updated channel environment and location. The effect of mobility and changing channel conditions, on the system, can be observed as follows:

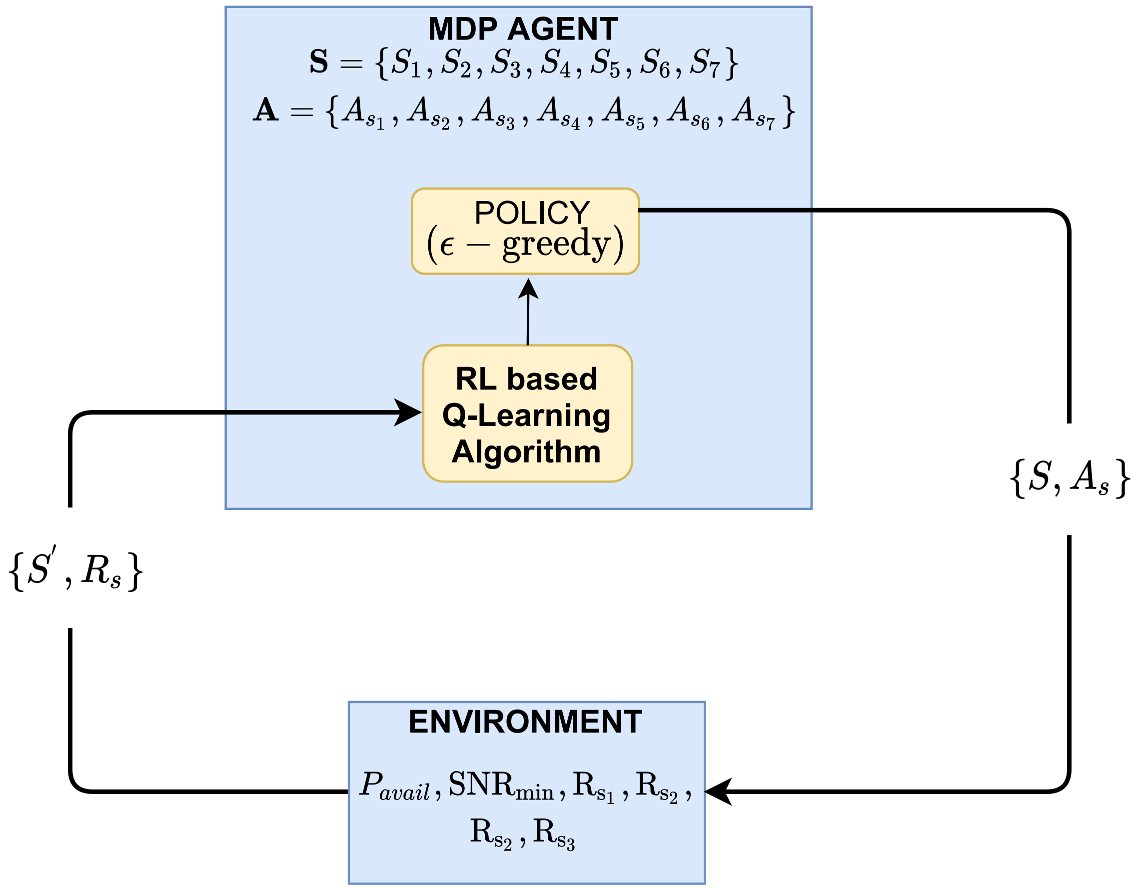



Figure 6.7: A high level illustration of the finite MDP reinforcement learning model to solve the reformulated problem.

$$T_{rd} = \frac{\mathrm{F_{rd}}}{\mathcal{D}_{rd}}, \tag{6.8}$$

$$\mathcal{D}_{rd} = \mathrm{BW} \log_2(1 + \mathrm{SNR_{rd}}), \tag{6.9}$$

$$\mathrm{SNR_{rd}} = \frac{\mathrm{P_{rd} h_{rd}}}{\mathrm{noise}}, \tag{6.10}$$

$$h_{rd} = \frac{c}{4\pi f_{rd} dist_{rd}}, \tag{6.11}$$

where $c$ is the speed of light constant, $BW$ is the bandwidth, $h_{rd}$ is the path loss function, $dist_{rd}$ denotes the parameter getting affected with the mobility of RN.

As RN changes its location, the distance between RN and its destined node needs to be updated which is being handled before every episode. The parameters $dist_{rd}$ and $noise$ directly affect the selection of suitable $\mathcal{D}_{rd}$ and $f_{rd}$. Once the RN learns its environment from running the episodes multiple times, Eq. (6.7) converges to the optimal decisions of $\mathcal{D}_{rd}$ and $f_{rd}$ for dynamic mobile environment.

After getting trained by Algorithm 5, the Q-tables of each state present the optimum values. These are used by Algorithm 6. Through the steps of Algorithm 6, it can be seen that once the packet is received, the RN is ready to take its decisions from the optimum values present for its states unlike the back-end algorithm where the system first has to perform an entire search to reach to the decisions. In Algorithm 6, once the packet is received, the RN just has to exploit the converged optimum values for taking optimum decisions.

**Algorithm 5** Backend Q-learning algorithm.

**Input:** $\alpha$, $\gamma$, $\epsilon$, episodes

**Output:** $Q^*(S_1, A)$, $Q^*(S_2, A)$, $Q^*(S_3, A)$, $Q^*(S_4, A)$

**for** $S$=$S_1, S_2, S_3, S_4$ **do**

  **for** *all actions* **do**

    *Initialize* $Q(S, A)$ *with any value*

**for** $S$=$S_5, S_6, S_7$ **do**

  **for** *all actions* **do**

    *Initialize* $Q(S, A)$ *of terminal states with zero*

**for** *each episode* **do**

  *Initialize* $S = S_1$

  *Use current location for* $dist_{rd}$

  *Use current channel situation for sensing channels and SNR estimation*

  **for** *each step of episode* **do**

    *Choose action from available action set with* $\epsilon$*-greedy policy*

    *Send to environment*

    *Observe corresponding reward (R) and next state (S')*

    *Update* $Q_(S, A) = Q(S, A) + \alpha[r + \gamma \underset{A}{max} Q(S', A) - Q(S, A)]$

    *ReturnS'*

  *Return once S==terminal state*

**Algorithm 6** Current time optimum Q-Value ($Q^*$) exploiting procedure for optimal multi-band channel allocation to RN.

**Input:** $F_{sr}$, $t_{ir}$, $\mathcal{D}_{sr}$, $\mathrm{SNR}_{\mathrm{sr}}$ buffsize, $f_{sr}$, $dist_{rd}$, $\mathrm{SNR}_{\mathrm{min}}$, $P_{avail}$

**Output:** $t_{out}$,$\mathcal{D}_{rd}$, $f_{rd}$

Save $t_{ir}$, the instant of receiving $F_{sr}$

**for** $S = S_1, S_2, S_3, S_4$ **do**

  Get current value of $maxQ(S, A)$ from back-end learning machine (Algorithm 5)

Select $f_{rd}$ from action space of $S_1$ having $maxQ(S_1, A) == Q(S_1, f_{rd})$

Get $S'$ corresponding to action $f_{rd}$

Select $\mathcal{D}_{rd}$ from action space of $S'$ having $maxQ(S', A) == Q(S', \mathcal{D}_{rd})$

**if** $S' == S_2$ **then**

  $t_{out} = t_{ir} + T_{sr} - T_{rd}$

**elseif** $S' == S_3$ **then**

  $t_{out} = t_{ir} + T_h$

**elseif** $S' == S_4$ **then**

  $t_{out} = t_{ir} + T_{sr}$

**return** $t_{out}$

Table 6.2: Considered simulation parameters.

| **Parameters** | **Values** |
|---|---|
| $F_{sr}$,$F_{rd}$ | 3000 (bits) |
| BW | 20 MHz |
| $\mathrm{SNR}_{\mathrm{min}}$ | -5 dB |
| Standard data-rate set | 6, 9, 12, 18, 24, 36, 48, 54 (Mbps) |
| Number of RNs | 4 nodes |
| buffsize | [400, 600, 200, 800] (bits) |
| $P_{avail}$ | [1, 3, 5, 4] (Watts) |
| Number of episodes | 2000 |
| $\epsilon, \gamma, \alpha$ | 0.1, 0.5, 0.1 |

### 6.4.1 Computational Complexity

There are two categories of the proposed RL-based approach. One is the front-end code given in Algorithm 6 and the other one is the back-end learning code shown in Algorithm 5.

It can be seen from Algorithm 5, there are three double nested FOR loops. Generally its computational complexity can be written as $3O(n^2) = O(n^2)$. To be specific the first double nested FOR loop has computational complexity of $O(4 \times I)$ where, $I$ represents the number of elements in the set: $\boldsymbol{A_{s_1}} \cup \boldsymbol{A_{s_2}} \cup \boldsymbol{A_{s_3}} \cup \boldsymbol{A_{s_4}}$. The second double nested FOR loop has computational complexity of $O(3 \times J)$, where $J$ represents the number of elements in the set: $\boldsymbol{A_{s_5}} \cup \boldsymbol{A_{s_6}} \cup \boldsymbol{A_{s_7}}$. The third double nested FOR loop has computational complexity of $O(ep \times 2)$, where $ep$ is the number of episodes and 2 represents two levels of state machine, initial and intermediate as, at terminal state the inner FOR loop exits.

In Algorithm 6, there is only one FOR loop but it runs for two times only: one for the initial state and then second time for the chosen next intermediate state. Hence its computational complexity can be written as O(2).

## 6.5 Performance Evaluation

In this section, the proposed reinforcement learning approach is evaluated based on computer-based simulations. First, the proposal is compared with conventional schemes of DF, random, and greedy channel selection [33] algorithms. Then, the effect of varying Q-learning parameters $\alpha$ and $\gamma$ on the learning process is provided. MATLAB scripts [43] are used to construct simulations using the parameters given in Table 6.2. The starting data rate from the main SN is set as $\mathcal{D}_{sr}$ = 18 Mbps with $\mathrm{SNR}_{\mathrm{sr}}$ = 8dB for plotting the results.

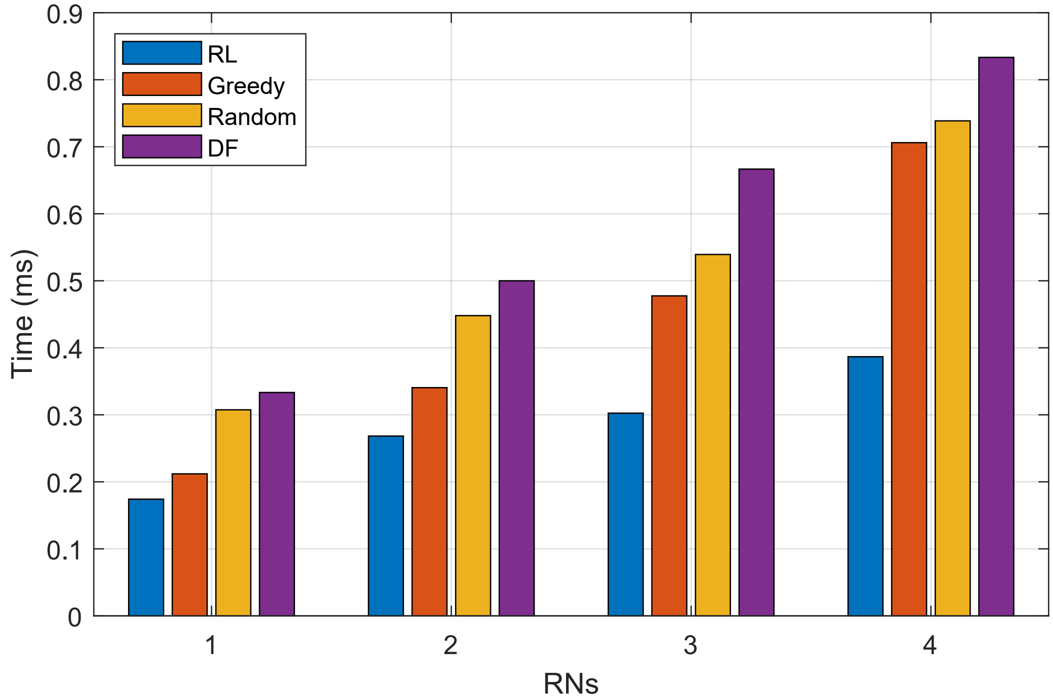



Figure 6.8: Comparison of the total transmission time for the considered methods with a varying number of RNs.

First, in Fig. 6.8, the proposed Q-Learning based reinforcement learning method is compared with the conventional DF scheme, the random channel assignment, and the greedy-heuristic algorithm [11] for the worst case scenarios where only low SNR channels are available with respect to the incoming packet's channel. It can be noticed from the results that using the conventional DF scheme (without multi-band), the total transmission time delay is the longest because of RNs receiving and transmitting the data packets one at a time.

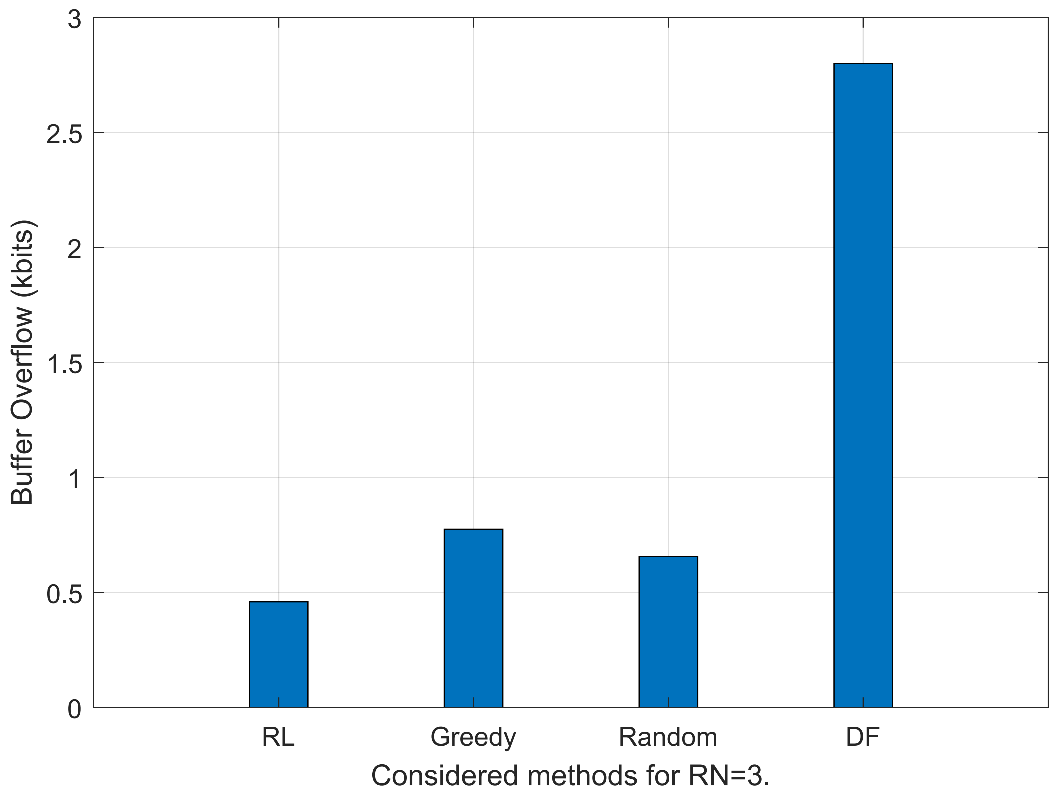



Figure 6.9: Buffer overflow comparison of RN=3.

In the random channel allocation, the RNs arbitrarily select channels, resulting in an unfeasible channel selection, which results in the second longest time to relay the data packets to DN. Furthermore, it can be observed that the greedy algorithm performs somewhat better than DF and random channel selection. However, in the worst case scenario, the proposed reinforcement learning algorithm provides more feasible, converged results.

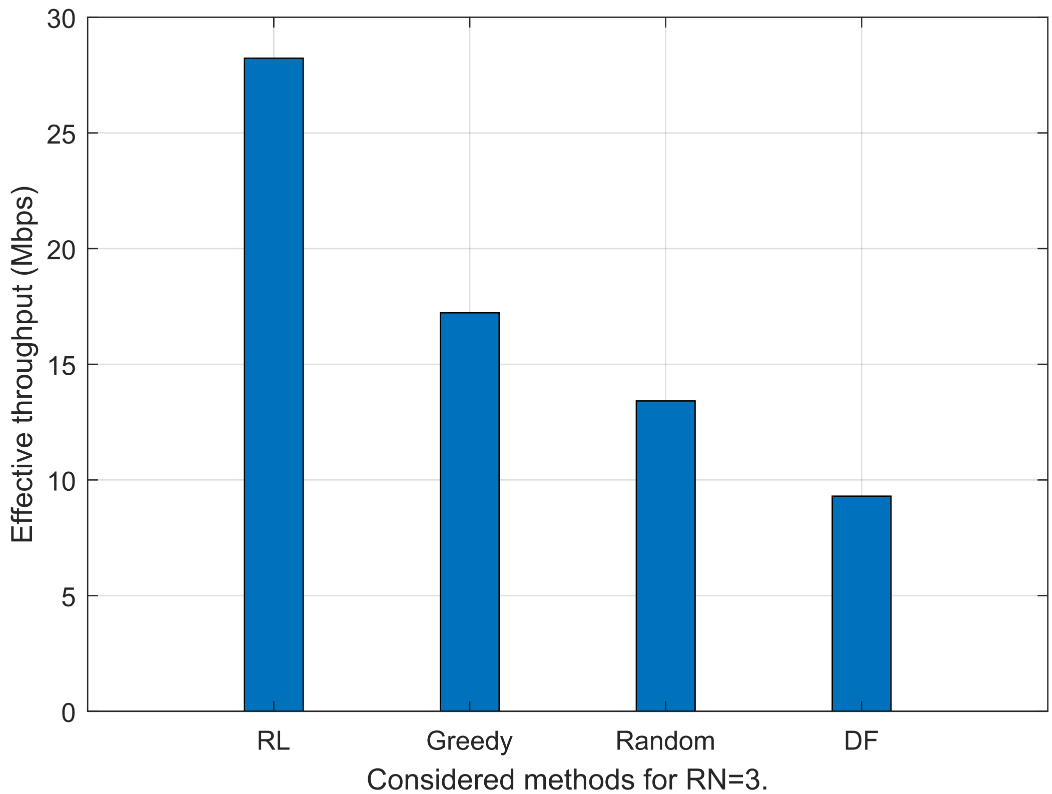



Figure 6.10: Effective throughput comparison of RN=3.

Next, in Fig. 6.9, the buffer overflow results of the proposal are provided, in contrast with the conventional DF scheme, the random channel assignment, and the greedy-heuristic algorithm. It can be seen from the results that even in the worst-case scenario, the proposed learning model still converges so as to induce a minimum number of buffer overflow bits as compared with all other approaches. Also, note that the buffer overflow of DF scheme is the highest which corroborates the need for a multi-band channel selection method at the distributed (RN) level.

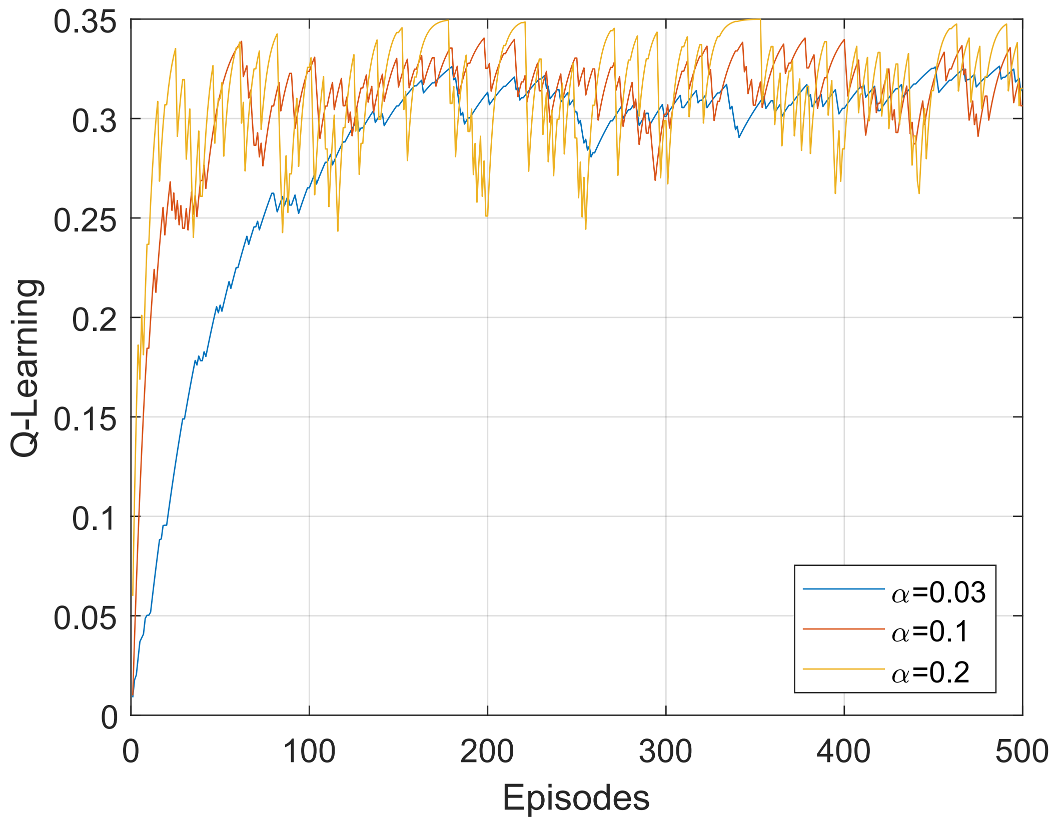



Figure 6.11: Effect of varying $\alpha$ with constant $\gamma = 0.1$.

Next, the effective throughput performance of the compared methods is observed. The effective throughput is defined as the number of successful packets received during total time of relaying. It can be seen from Fig. 6.10 that the effective throughput of the proposed reinforcement learning approach is the highest due to its ability to maximize the time efficiency while minimizing buffer overflow. On the other hand, the DF method suffers from the least effective throughput by exhibiting an exponential rise in the packet latency.

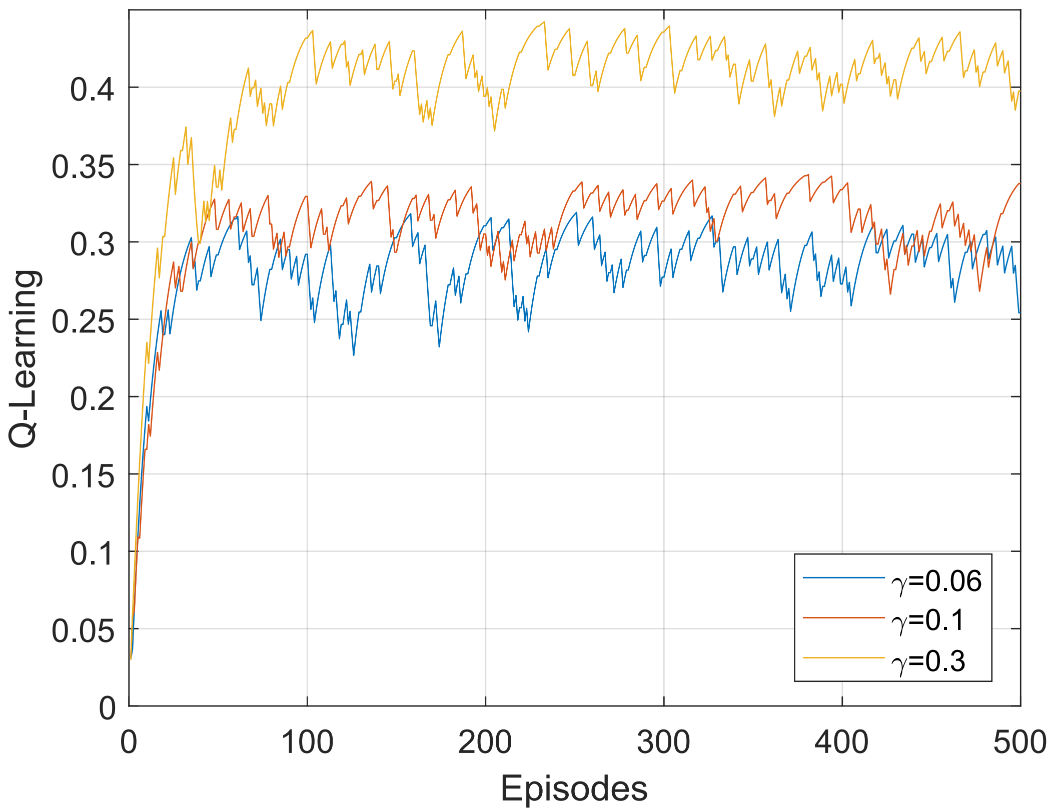



Figure 6.12: Effect of varying $\gamma$ with constant $\alpha = 0.1$.

Fig. 6.11 demonstrates the effect of varying the learning rate, $\alpha$, of the proposed reinforcement learning algorithm while keeping the future reward parameter, $\gamma$, as a constant. If the learning rate is too fast, the algorithm converges quickly; but it does not maintain a stable value for long. On the other hand, as the learning rate is made smaller, the algorithm takes longer to reach to the convergence; but it gets stable faster.

Finally, Fig. 6.12 demonstrates the effect of varying the future reward parameter, $\gamma$, of the proposed reinforcement learning algorithm while setting the learning rate, $\alpha$, as a constant. When the future rewards are given lesser values, the algorithm struggles to reach the optimum value. On the other hand, the more the value given to the future rewards, the better the algorithm reaches to its optimum value.

## 6.6 Conclusion

Relay-based networks have been increasingly adopted in a wide range of scenarios, from device-to-device networks to drone-cells for improved capacity and coverage. In 5G+ and 6G integrated networks, the role of the relay-based networks will be different from their predecessors due to various reasons. A key reason is the introduction and incorporation of various frequency bands and their simultaneous use by source, destination, and relay nodes.

While high frequency bands provide much higher capacity, they are constrained with more stringent path loss and blocking; and the channel conditions may drastically change in ultra-high frequency spectra in the upper GHz and THz level, such as visual light communication (VLC). Optimal channel allocation to these relay nodes to combat the real-time traffic load variation and other network dynamics including user mobility is shown to be a computationally hard problem, attempting to solve which even for a low number of relay nodes requires a centralized oracle-like platform (e.g., a software defined network controller, or a central cloud server) to compute optimal channel allocation decisions.

This is intractable for wireless relay nodes, which cannot wait for receiving a central decision while the network dynamics continue to change. This warrants a smart and distributed solution that needs to native to the relay nodes.

In this chapter, this complex problem is addressed as an MDP, optimal solution of which is also shown to be expensive. This motivated us to customize a reinforcement-learning method to solve the problem with near-optimal performance in real-time at the relay nodes.

Convergence shows the fast-learning curve for the proposal. Comparative results also demonstrate that the proposed reinforcement learning-based approach achieves comparable performance to that of the centralized benchmark and also outperforms several existing techniques in terms of packet transmission time, buffer overflow, and effective throughput.

## 6.7 Publications Resulted from This Chapter

- B. Mughal, Z. M. Fadlullah, M. M. Fouda and S. Ikki, "Optimizing Packet Forwarding Performance in Multi-Band Relay Networks via Customized Reinforcement Learning," Submitted to *IEEE Transactions on Vehicular Technology*, Dec 2021.

# Chapter 7

# Conclusion and Future Works

This chapter consists of two main sections: The first section, "Conclusion" discusses summary of this thesis along with deduction from proposed work. The second section discuss possible future work on the topic.

## 7.1 Conclusion

With technology being excelled towards 5G+, increasing coverage is not just some option to be availed, it is in fact a requirement. Hence presence of intermediate nodes connected via relay topology is one essential part of any network. Such privilege of relaying, has added data latency which straight forwardly means low data rate and/or distorted reception.

In this thesis, the capacity versus coverage trade-off has been taken into account for next generation communication networks. Along which, the shortcoming of existing relay-assisted techniques to improve the service coverage is also discussed. The research gap in the existing literature is identified for using multiple bands in RNs simultaneously without considering the impact of delay.

### 7.1.1 Greedy

This issue is then addressed by formulating an optimal centralized heterogeneous band/channel allocation method for the RNs to achieve minimum latency. This centralized approach results in adding more data latency of packet re-forwarding because of its non-deterministic execution nature.

The centralized approach being not a practical option for deployment of resource-constrained RNs such as D2D networks, a heuristic-based distributed channel allocation algorithm [11] is proposed with the local information available at individual RNs.

Conducted simulation results demonstrated that our proposed distributed approach outperformed the conventional DF method, and also scaled well with the increasing number of RNs in contrast with the centralized method to solve the formulated optimization problem.

### 7.1.2 Game

Due to greedy nature of precisely proposed distribution solution, there is still room for efficiency improvement. Hence the research on proposing greedy-based heuristic is further extended, by proposing a better online game theory-based distributed model [41].

This proposed game model is then compared with its corresponding bipartite graph-based oracle method. The simulations are then conducted as compared with greedy and other conventional schemes to prove the efficiency of proposed distributed sequential game model.

### 7.1.3 Nash

In the game theory, in order to find the best strategy for each player, theorem of Nash Equilibrium is used (1.5(3.)). For sequential games with perfect information, in order to find the best strategy for each player, sub-game perfect equilibrium is found using backward induction. This thesis also provides the complexity analysis of proposed game-theoretic model, with simple yet illustrative examples, based on which it is deduced that the sequential game can be converged and stabilized by obtaining the Nash Equilibria of the sub-games within the main game.

### 7.1.4 Reinforcement Learning

In order to handle certain mobility of user equipment and the stochastic channel environments, this thesis also present an ultimate solution to the problem based on artificial intelligence algorithm-reinforcement learning (1.5(4.)). The simulation results show the practicality and the enhanced performance of proposed reinforcement learning algorithm which is solved using Q-Learning geared with $\epsilon$-greedy.

## 7.2 Future Work

This section provides the possibility of future work on the topic so as to enhance the present work. In addition to adjusting parameters in this thesis, work can be done on adding option of available intermediate RNs queue, the MDP state machine design, the MDP rewards re-designing and implementing this work considering the flying devices.

### 7.2.1 RNs queue selection

In this thesis, single pre-selected queue of RNs is considered. Once a SN transmits data packet, the RN and/or the rest of the queue of participating intermediate RNs is already selected and known, as per the assumption of this work, but as a consequence SN has only one path (queue) to rely on.

In future, work can be done in this aspect of the thesis by assuming multiple available paths (queues) of participating intermediate RNs. It can be designed in such a way that each queue trains itself for finding its feasible total packet forwarding time till destination device. Then the queue with the least total time is chosen for actual packet forwarding, starting from SN via chosen RN queue till DN.

Such set up might increase the system overheads but if Reinforcement Learning approach is used, this complicated setup can explore more options and providing better results. Work done on this aspect of thesis, can increase the probability of achieving even lesser total packet forwarding time.

### 7.2.2 MDP States

Currently, proposed MDP consists of 7-state model focusing on the states where, once the packet is received at RN, SS and SNR estimation is being done together, leading towards availability of either no channels available or lesser SNR channels available or greater SNR channels available, for packet re-forwarding. These states then get at next-states of either enough power not available, or QoS or feasible channel.

In future work can be done on comparing this proposed model by designing single state setup by having time instant as a next state. That is, say, at one time instant if some channel is being chosen, then at next time instant some other channel would be chosen (with all processing happening at the back-end) then a future estimate can be found as a continuous learning process.

Alternatively, another comparison can also be done by designing more expanded states. That is, say having 2 separate states for SS and SNR estimation. Such kind of expanded state design could It is compelling to see comparison among single state, currently proposed 7-state model and future expanded state design.

### 7.2.3 MDP Rewards design

MDP rewards designing is another aspect of the proposed work that can be further explored. In proposed 7-state model as per figs. 6.6 and 6.5, rewards are designed mainly, on the basis of time taken by the packet. In future, these rewards can be re-designed by using many other features of data transmission like power, QoS, location and/or distance between the nodes. It would be indeed compelling to see the comparison between results achieved using current reward design and the results obtained from newly designed rewards.

### 7.2.4 Flying RNs

Recently, drones, also known as Unmanned Aerial Vehicles (UAVs), Network Flying Platforms (NFPs) and Low Altitude aerial Platforms (LAPs), have gained notable interest in research and in various commercial, military, civilian, agricultural, and environmental applications. The examples of such wide range of applications are: wireless coverage, relay for adhoc networks, precision agriculture, disaster monitoring, managing wildfire, wind estimation, aerial border surveillance, traffic monitoring, remote sensing, power lines monitoring, construction, remote blood delivery, search and destroy operations, etc.

Currently, proposed model considers ground mobile devices only. In future, the proposed MDP Reinforcement learning model can be analyzed using NFP as one of the participating intermediate RNs. By using NFPs as an option, parameters like altitude and trajectory can also be made part of states and rewards re-designing. It will be another exciting aspect of the system that can be explored and then compared with the system dependant on ground devices only.

# References


[1] X. Bu *et al.*, "Optimization on cooperative communications based on network coding in multi-hop wireless networks," in *International Wireless Communications and Mobile Computing (IWCMC)*, pp. 384–387, 2020.

[2] A. Ben-Bassat, S. Gross, A. Nazimov, A. Ravi, B. Khamaisi, E. Banin, E. Borokhovich, N. Kimiagarov, P. Skliar, R. Banin, S. Zur, S. Reinhold, S. Bruker, T. Maimon, U. Parker, and O. Degani, "10.5 a fully integrated 27dbm dual-band all-digital polar transmitter supporting 160mhz for wifi 6 applications," in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, pp. 180–182, 2020.

[3] L. F. Abanto-Leon, A. Koppelaar, and S. H. de Groot, "Graph-based resource allocation with conflict avoidance for V2V broadcast communications," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017.

[4] H. Azegami, *Basics of Optimization Theory*, pp. 45–103. Singapore: Springer Singapore, 2020.

[5] A. McEachern and D. Ashlock. 2017.

[6] J. Chen, "Nash equilibrium." https://www.investopedia.com/terms/n/nash-equilibrium.asp. Accessed on: July 2021.

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

[8] L. Zhang, J. Sun, Q. Zheng, L. Chen, and Z. Wang, "Frame structure design method for large carrier frequency offset estimation under low snr in burst transmission," in *2019 IEEE 5th International Conference on Computer and Communications (ICCC)*, pp. 865–870, 2019.

[9] Y. Mao, M.-L. Zhu, T. Sun, Y.-Y. Dong, and C.-X. Dong, "Automatic modulation classification based on snr estimation via two-stage convolutional neural networks," in *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 294–298, 2021.

[10] B. K. Das, "Low snr spatially sparse channel estimation in millimeterwave hybrid mimo systems based on adaptive filtering framework," in *2020 IEEE Applied Signal Processing Conference (ASPCON)*, pp. 143–146, 2020.

[11] B. Mughal, S. Hussain, and A. Ghafoor, "Cooperative subcarrier sensing using antenna diversity based weighted virtual sub clustering," *China Communications*, vol. 13, no. 10, pp. 44–57, 2016.

[12] M. Li *et al.*, "Optimal scheduling for multi-radio multi-channel multi-hop cognitive cellular networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 1, pp. 139–154, 2015.

[13] N. Egashira *et al.*, "Low latency relay processing scheme for WLAN systems employing multiband simultaneous transmission," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.

[14] N. Egashira *et al.*, "Experimental evaluation of adaptive simultaneous transmission timing control considering idle/busy probability for multi-band wireless lan," in *2018 Asia-Pacific Microwave Conference (APMC)*, pp. 881–883, 2018.

[15] A. Hanyu *et al.*, "Adaptive frequency band and channel selection for simultaneous receiving and sending in multiband communication," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 460–463, 2019.

[16] Z. M. Fadlullah, Y. Kawamoto, H. Nishiyama, N. Kato, N. Egashira, K. Yano, and T. Kumagai, "Multi-hop wireless transmission in multi-band wlan systems: Proposal and future perspective," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 108–113, 2019.

[17] L. Zhang *et al.*, "Multi-channel allocation algorithm based on congestion avoidance in wearable wireless sensor network," in *2018 27th Wireless and Optical Communication Conference (WOCC)*, 2018.

[18] K. Ding, H. Zhao, X. Hu, and J. Wei, "Distributed channel allocation and time slot optimization for green internet of things," *Sensors*, vol. 17, no. 11, 2017.

[19] N. Egashira, K. Yano, S. Tsukamoto, J. Webber, and T. Kumagai, "Low latency relay processing scheme for wlan systems employing multiband simultaneous transmission," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.

[20] L. Abdelwedoud, A. Busson, and I. G. Lassous, "Use of a weighted conflict graph in the channel selection operation for Wi-Fi networks," in *2021 16th Annual Conference on Wireless On-demand Network Systems and Services Conference (WONS)*, pp. 1–4, 2021.

[21] H. Zhang, X. He, Q. Wu, and H. Dai, "Spectral graph theory based resource allocation for IRS-assisted multi-hop edge computing," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, 2021.

[22] Y. Liu, Z. Wen, N. C. Beaulieu, D. Liu, and X. Liu, "Power allocation for swipt in full-duplex af relay interference channels using game theory," *IEEE Communications Letters*, vol. 24, no. 3, pp. 608–611, 2020.

[23] A. Khodmi, S. B. Rejeb, N. Agoulmine, and Z. Choukair, "Joint user-channel assignment and power allocation for non-orthogonal multiple access in a 5G heterogeneous ultra-dense networks," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1879–1884, 2020.

[24] M. Di Felice, L. Bedogni, and L. Bononi, *Reinforcement Learning-Based Spectrum Management for Cognitive Radio Networks: A Literature Review and Case Study*, pp. 1–38. Singapore: Springer Singapore, 2018.

[25] X. Huang, T. Han, and N. Ansari, "On green energy powered cognitive radio networks," *IEEE Communications Surveys and Tutorials*, vol. 17, 04 2014.

[26] M. S. Parwez and H. Olufowobi, "Cost-constrained handoff in next generation heterogeneous wireless networks," in *2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pp. 911–916, 2018.

[27] K. Chen, Y. Wang, Z. Fei, and X. Wang, "Power limited ultra-reliable and low-latency communication in uav-enabled iot networks," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2020.

[28] "Documentation: Mixed-integer linear programming algorithms." https://www.mathworks.com/help/optim/ug/mixed-integer-linear-programming-algorithms.html. Accessed on: Aug 2020.

[29] "Documentation: Mapping 802.11 standards to wlan toolbox configuration objects." https://www.mathworks.com/help/wlan/ug/mapping-802-11-standards-to-wlan-toolbox-configuration-objects.html. Accessed on: Aug 2020.

[30] W. Wang, Y. Shen, and Y. Wang, "Low-complexity non-data-aided snr estimation for multilevel constellations," *IEEE Communications Letters*, vol. 24, no. 1, pp. 113–116, 2020.

[31] "Documentation: wlannonhtconfig." https://www.mathworks.com/help/releases/R2016a/wlan/ref/wlannonhtconfig.html. Accessed on: Aug 2020.

[32] "Game theory." https://en.wikipedia.org/wiki/Game_theory. Accessed on: July 2021.

[33] B. Mughal, Z. M. Fadlullah, and S. Ikki, "Centralized versus heuristic-based distributed channel allocation to minimize packet transmission delay for multiband relay networks," *IEEE Networking Letters*, vol. 2, no. 4, pp. 180–184, 2020.

[34] C. E. Shannon, *Programming a Computer for Playing Chess*. Springer New York, 1988.

[35] F. Forgó, "John von neumann's contribution to modern game theory," *Acta Oeconomica - ACTA OECON*, vol. 54, pp. 73–84, 05 2004.

[36] P. G. Spirakis, "Extensive games with perfect information: Stackelberg's duopoly game." https://cgi.csc.liv.ac.uk/~spirakis/COMP323-Fall2015/lectures/week05.pdf. Accessed on: July 2021.

[37] M. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2000.

[38] J. D. Morrow, *Game Theory for Political Scientists*. Princeton University Press, 1994. Accessed on: July 2021.

[39] B. Mughal, Z. Fadlullah, M. M. Fouda, and S. Ikki, "Allocation schemes for relay communications: A multi-band multi-channel approach using game theory," *IEEE Sensors Letters*, pp. 1–1, 2021.

[40] Z. M. Fadlullah, C. Wei, Z. Shi, and N. Kato, "Gt-qosec: A game-theoretic joint optimization of qos and security for differentiated services in next generation heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1037–1050, 2017.

[41] B. Mughal, S. Hussain, and A. Ghafoor, "Cluster-based cooperative sub-carrier sensing using antenna diversity-based weighted data fusion," *Arabian Journal for Science and Engineering*, vol. 41, no. 9, pp. 3425–3439, 2016.

[42] J. Haghighat and W. Hamouda, "Decode-compress-and-forward with selective-cooperation for relay networks," *IEEE Communications Letters*, vol. 16, no. 3, pp. 378–381, 2012.

[43] "Matlab." https://www.mathworks.com. Accessed on: Dec 2021.