RESOURCE ALLOCATION MECHANISM FOR MULTIUSER MULTICHANNEL COGNITIVE RADIO NETWORKS

by

Shi Wang

Submitted in partial fulfillment of the requirements for the degree

Philosophiae Doctor (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

March 2018

SUMMARY

RESOURCE ALLOCATION MECHANISM FOR MULTIUSER MULTICHANNEL COGNITIVE RADIO NETWORKS

by

Shi Wang

Promoter(s):	Prof. B. T. Maharaj
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Philosophiae Doctor (Electronic Engineering)

Keywords:	Cognitive radio, channel allocation protocol, adaptive modulation and coding, distribution probability matrix, maximize throughput, queueing theory, queueing analytical framework, heuristics, optimization

Cognitive radio (CR) technology has been considered one of the most promising solutions to make the best potential use of the scarce spectrum resources in next generation wireless communication. Because of the stringent limitations on the utilization of the radio resources in a cognitive radio network (CRN), the available resources are usually scarce and dynamic. Developing appropriate mechanism to construct allocation protocols for CRN that can assign the resources in a fair and efficient manner to diverse users with different features is essential to achieve the promise of the CR technology in reality and worth studying. "The mechanism for resource allocation protocol design in CRN" describes the topic of this research work and the contents of this thesis are organized around it.

Before investigating how to design appropriate mechanism for CR resource allocation (RA) protocols, the model of the radio resource and the model of how a CR system utilizes the resource have to be clarified. In this thesis, by a comprehensive study of the relevant literatures, the modeling of critical components of a CR system and the theory behind it are identified and investigated. The existing protocols are also analyzed to address the limitations and gaps in the design of allocation protocols.

An important feature, if not the most, that the allocation protocol for the CRN should possess is the ability to adjust the strategy of RA when the circumstances that affect the data transmission of the CR system varies, such as the activity of the license users or the variety of CR users' demands. According to such considerations, the mechanism, namely distribution probability matrix (DPM), is designed to guide and describe the channel allocation protocol quantitatively. To testify how this mechanism works, the protocols based on the DPM, together with reference protocols in a multichannel CRN, are analyzed by a queuing model. The numerical results shows the flexibility and adaptability that the protocols based on the DPM can achieve.

Another important function of the DPM mechanism is the potential to describe complex protocols with specific principles and to help design protocols aimed at specific objectives. To demonstrate this, a maximum throughput (MT) protocol that is aimed at maximizing the overall throughput of the CR system in a multi-user multi-channel scenario is structured using the DPM mechanism. The existing reference protocols, including maximum rate and random allocation, are described under the DPM mechanism as well as the MT protocol. All the protocols are implemented in a multiuser multichannel CRN and analyzed by a queueing analytical framework proposed by the author that is capable of acquiring the performance metrics of each CR user independently with efficiency. Numerical results show that the MT protocol is able to outperform the existing protocol in terms of the overall throughput of the CRN.

The last problem identified and addressed in this thesis is to investigate the potential of specific performance metrics that a CR system can achieve under certain conditions with the help of the DPM mechanism. An optimization problem that is aimed at maximizing the overall weighted throughput of a multi-user multi-channel CRN under the delay constraint is formulated and solved using the DPM mechanism. The key factors that affect the optimal weighted throughput and the optimal allocations are investigated through the numerical results obtained by the proposed queueing analytical framework. The advantages of the DPM mechanism in optimizing the performance of CRN is also revealed by the formulation and solution of the optimization problem.

The RA protocol is the essential kernel to address the RA problem in the CRN context. This thesis, by providing an objective and effective mechanism that works as a universal tool to describe and evaluate the RA protocols, succeeds in making effective contribution to addressing the RA problem in the CRN and further, helping the CR technology realize its meaningful promises.

LIST OF ABBREVIATIONS

ACK	Acknowledgment
AMC	Adaptive modulation and coding
ARQ	Automatic repeat request
BER	Bit error rate
BPI	Block probability index
BPSK	Binary phase-shift keying
BW	Bandwidth
CR	Cognitive radio
CRN	Cognitive radio networks
CSI	Channel state information
DPM	Distribution probability matrix
DTMC	Discrete time Markov chain
FSMC	Finite state Markov chain
GDP	Gross domestic product
ICT	Information communications technology
IEEE	Institute of Electrical and Electronics Engineers
MR	Maximum rate
MT	Maximum throughput
NYC	New York City
OP	Overtime probability
PDF	Probability density function
PHY	Physical
PU	Primary user
QAM	Quadrature amplitude modulation

QoS	Quality of service
QPSK	Quadrature phase-shift keying
RA	Resource allocation
RF	Radio frequency
RR	Round robin
SMR	Surface movement radar
SNR	Signal-to-noise ratio
SU	Secondary user

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 PROBLEM STATEMENT

1.1.1 Context of the problem

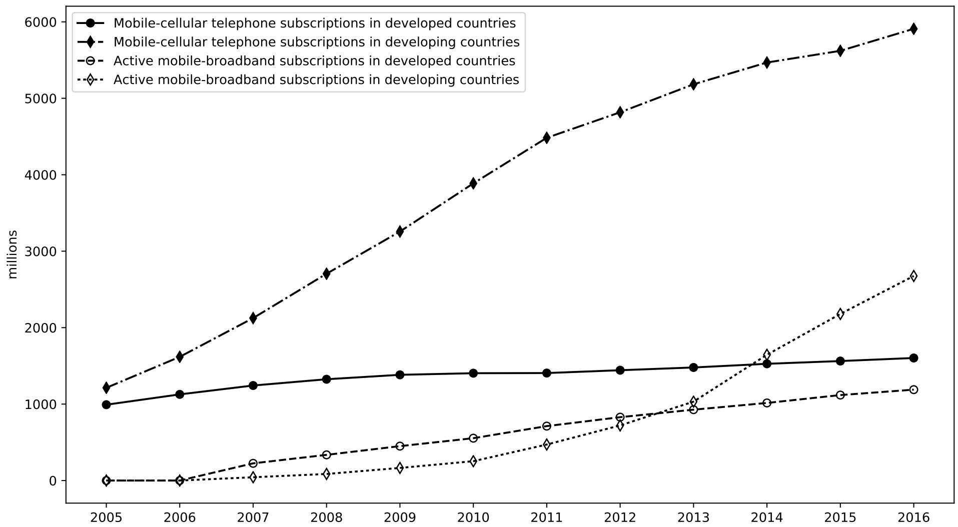
The enormous development of wireless communication technology has occurred in the last decade and changed everyone's life worldwide. Make the mobile communication as an example: Figure 1.1 shows the number of mobile-cellular telephone subscriptions and mobile-broadband subscriptions in developed and developing countries from 2005 to 2016 [1].

The development of wireless communication not only makes people's life more convenient but also promote economic prosperity. According to the World Bank database [2], the information communications technology (ICT) sector contributes 6% of total gross domestic product (GDP) all over the world and the contribution to GDP growth is significant, as shown in Table 1.1.

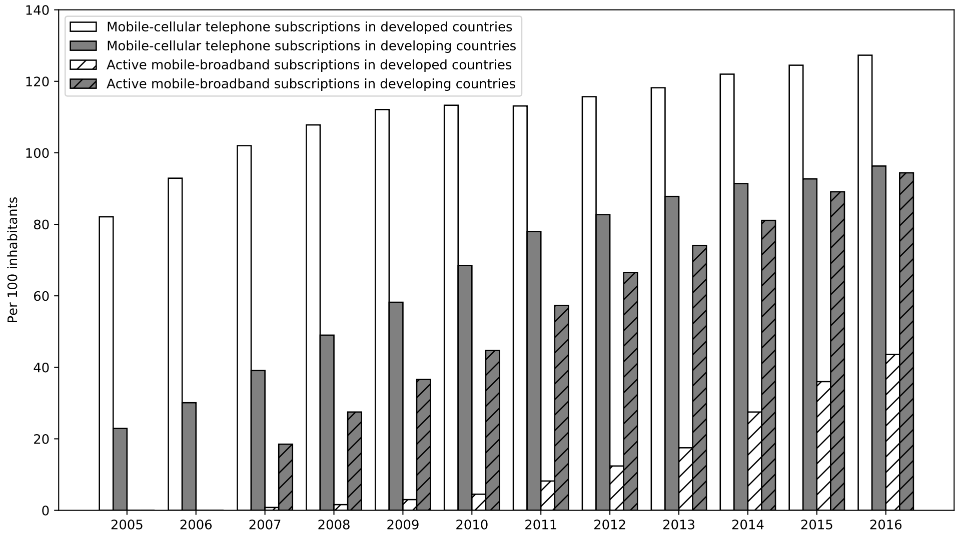
Table 1.1. The size of the ICT sector and its contribution to GDP growth.

		Year			
		1995	2000	2005	2010
39 developed countries	GDP growth	3.1%	2.5%	1.1%	1.8%
	ICT share	24%	23%	38%	22%
91 developing countries	GDP growth	3.4%	6.1%	6.6%	5.6%
	ICT share	13%	10%	19%	15%

Since the demand for wireless communication, especially broadband communication, is rapidly increasing, radio spectrum is becoming a scarce and precious resource. The traditional spectrum



(a) Number of mobile-cellular telephone subscription and mobile-broadband subscription.



(b) Coverage of Mobile-cellular telephone subscription and mobile-broadband subscription.

Figure 1.1. Mobile-cellular telephone subscription and mobile-broadband subscription.

allocation policy is fixed allocation, which means only the licensed user is allowed to use the assigned spectrum. This is successful in avoiding interference between different users when the spectrum resource is sufficient. However, with the demand for spectrum increasing, it becomes necessary to reconsider this policy. Several spectrum occupancy measurements have been conducted worldwide [3–8]. The results show that the average efficiency of the usage of the spectrum is at a low level in the relevant countries. For example, Table 1.2 shows the results of spectrum occupancy measurement in New York City (NYC) and Chicago [8]. Moreover, the situation of spectrum occupancy also varies from time to time during the day, day to day in a week [9].

Table 1.2. Summary of spectrum occupancy in each band in New York City and Chicago

Start Freq (MHz)	Stop Freq (MHz)	Bandwidth (MHz)	Application	NYC average percent occupied	Chicago average percent occupied
806	902	96	Cell phone and SMR	46.3%	54.8%
1240	1300	60	Amateur	0%	0%
1850	1990	140	PCS, Asyn, Iso	33.8%	42.9%

The underutilization of the spectrum makes it necessary to think about the utilization of the unused spectrum at a particular time and in a specific geographic location. This has led to the invention of the concept of CR in [10]. Following is a definition of CR given by Simon Haykin in [11]:

> Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming radio frequency (RF) stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind:
>
> - highly reliable communications whenever and wherever needed;
> - efficient utilization of the radio spectrum.

The original licensed user of the spectrum that a CR system tries to utilize is usually named primary user [12–14]. The user that a CR system serves is named secondary user.

Various potential applications of CR have been explored in many fields, such as smart grids [14–16], Internet of things [17–19], Internet of vehicles [20–22]. CR is also a promising technology for 5G cellular networks to tackle the challenges, such as highly distributed mobile traffic across space and time and increasingly diverse mobile traffic [23]. There is a published standard for CR: IEEE 802.22 [24–26].

1.1.2 Research gap

Form the investigation on the spectrum occupancy, such as in [9], it is evident that the radio resource available to the CR system is dynamical. Hence, the CR system must consider the time varying feature of the radio resource before it allocates the resources to its users. This requires the channel allocation protocol of the CR system to be flexible and adaptable to any possible situations to make full use of the radio resources. Because the CR users are not licensed to use the spectrum, the demand of communication cannot be guaranteed to be fulfilled as the licensed user, also as known as PU. This complicates the management of the radio resource of the CR system further. Firstly, a quality of service (QoS) evaluation standard different from the traditional communication system is required to describe the extent to which the CR users' demand can be fulfilled. Then a summary of the evaluations is needed to serve as guide in the radio resource allocation.

Most existing models in the literature on RA in cognitive radio networks (CRNs) do not treat the protocol in a dynamic way and thus neglect to allocate the resources taking into account the dynamic of the radio resource, the demand and achievable QoS of CR users.

1.2 RESEARCH OBJECTIVES

This research is aimed at studying the following subjects that are important issues for RA in CRN and have not yet been investigated well:

- Identifying a quantitative description of the dynamic radio resources available to the CR system taking into account PU activities and spectrum sensing process.
- Designing a novel radio RA mechanism with the following features:
  - The ability to direct the RA in a multi-user multi-channel CRN.
  - The ability to describe the result of the RA under various existing protocols and evaluate them in a multi-user multi-channel CRN.
  - Simplicity to configure and adaptability under various analytical frameworks.
- Developing a comprehensive framework that is capable of analyzing the relationship between the radio RA protocol and the performance in a multi-user multi-channel. This framework

provides various interfaces that can easily be configured by adjusting parameters to describe key factors that affect the performance of the CRN, such as the arrival process of CR users, the transmission conditions, the activity model of the PU and so on. The performance evaluations of every individual CR user can be revealed independently by this framework.

- Establishing a set of performance evaluation criteria to indicate the QoS of the CR users with the performance evaluations obtained using the framework. Using a combination of the criteria, identifying a practical optimization objective of radio RA and formulating a optimization problem with the channel allocation protocol settings as the input argument.
- Finding a practical method to obtain solution to the optimization problem mentioned above taking into account the concerns of both computing complexity and accuracy.

1.3 HYPOTHESIS, APPROACH AND RESEARCH GOALS

The hypothesis for this research can be stated as follows: “If the channel allocation protocol is modeled in a flexible and adaptable way and the CR system is described comprehensively by such protocol, the practical objective of CR can be described and the optimal resource allocation to achieve certain objectives can be obtained efficiently and accurately.” The validation of the hypothesis, which is proved by the statements in the following chapters of this thesis, provides a route to analyze RA problem and achieve better performance in the CRN from the view of the basic components to a larger scope because the proposed analytical model is expandable with standard interfaces of input and output parameters. This is a positive advancement in the field of RA in CRN.

The progress of this research work follows the well-established pattern for performing technical research in electronic engineering and telecommunication systems, particularly on wireless communication systems. In general, this research is divided into several stages and the details of each stage are as follows:

- **Literature survey:** The objective of this part of work is to investigate the analytical tools and theory used in relevant literature to study RA problem in CRN, to identify the research gaps and to determine the direction for the research work. The information gathered and the discoveries made during this stage were summarized in chapter two of the thesis.

- **System modeling:** The design of the system model structure and the input/output parameter patterns are the key issues in solving engineering problems. During the development of the system model to address the RA problem, the following steps are carried out in sequence.

  1. Identifying the omitted but important features in existing model and analyzing the reason why they have been omitted.
  2. Estimating the difficulties of implementation of the important features and determining the features to achieve in the model.
  3. Making adjustments to the design of the model considering the adaptability and expandability of the model.

  The details of the models and the way in which the challenges above are handled were presented in chapter three to five of the thesis.

- **Simulation and numerical analysis:** The system model was built and simulated in MATLAB software. The setting of the parameters of the model was determined before the running of the simulation considering both the generality and the computing complexity. The numerical results were organized by categories of input and output parameter entities. Then various relationships between different pairs of input and output parameter entities were investigated to reveal the performance of the new model under preset situations that were represented by a group of input parameters not included in the pair of input and output.
- **Verification and validation of results:** The verification and validation of the results was carried out as follows. First, similar relationships between the input and output parameters from related publications in the literature were obtained using this study's analytical framework in identical form. Then quantitative comparison was made to illustrate the advantages of the new model.
- **Thesis write-up:** The conclusion of the research work was delivered with this detailed documentation combined with various findings of the research and presented in this thesis.

1.4 RESEARCH CONTRIBUTIONS

In the research work presented in this thesis, the following contributions have been made to extend the boundaries of the knowledge on RA protocol in CRN:

- A novel channel allocation mechanism, namely the distribution probability matrix (DPM), is established with the following features. First, using this mechanism, a set of channel allocation protocols that is flexibly configurable and adaptable to various application scenarios can be constructed. Second, existing protocols can be described by this mechanism and evaluated efficiently and precisely. Third, the potential of the performance, such as throughput, of a multi-user multi-channel CRN can be revealed by this mechanism.
- An analytical framework based on the queueing theory is designed to model multi-user multi-channel CRN with the following features. First, various channel allocation protocols can be implemented in this framework with the help of the DPM mentioned above. Second, the performance metrics of the CR users can be obtained independently and efficiently, which makes the analysis of the overall performance of the multi-user multi-channel CRN simple and precise.
- A simplified method to analyze delay performance is proposed. The analysis of the exact delay performance is complicated, especially in a multi-user multi-channel CRN with a varying data transmission rate. This simplified method can estimate the delay performance approximately, but efficiently and furthermore, the investigation based on the delay performance, such as finding the optimal allocation schemes to minimize the delay, is made practicable.
- A heuristic solution to find the optimal allocation in a multi-user multi-channel CRN is developed based on the DPM. The feasible set of channel allocation results is structured by the DPM mechanism and a heuristic algorithm whose precisions can easily be adjusted according to the requirement is developed.

Full-length papers in peer-reviewed conference proceedings:

S. Wang, B. T. Maharaj, and A. S. Alfa, “Resource allocation and performance measures in multi-user multi-channel cognitive radio networks”, in Vehicular Technology Conference (VTC Spring), 2016, Nanjing, IEEE.

Peer-reviewed, ISI rated journal articles:

S. Wang, B. T. Maharaj, A. S. Alfa, “Queueing analysis of performance measures under configurable channel allocation in cognitive radio”. Accepted by IEEE Transactions on Vehicular Technology.

S. Wang, B. T. Maharaj, A. S. Alfa, “A maximum throughput channel allocation protocol in multi-channel multi-user cognitive radio network”. Accepted by Journal of Communications and Networks.

S. Wang, B. T. Maharaj, A. S. Alfa, “A Virtual Control Layer Resource Allocation Framework for Heterogeneous Cognitive Radio Network”. Submitted to IEEE Journal on Selected Areas in Communications, series on network softwarization & enablers.

1.5 OVERVIEW OF THESIS

The remainder of the thesis is structured as follows:

Chapter two presents detailed background knowledge related to the research in this thesis. The basics of how to model the radio resources is introduced, followed by a summary of existing RA protocols. Then the weakness of the principle of the protocols is revealed and possible directions that can improve the analysis of the CRN are suggested.

In Chapter three, the DPM mechanism for channel allocation protocol in CRN is introduced. The two basic components of the DPM mechanism, the allocation indicator matrix and the effective duration, are introduced, followed by the explanation of the typical usages of the DPM. In the next part of the chapter, the DPM is implemented as a channel allocation protocol in a multi-channel CRN. Then a queueing analysis is carried out to acquire the performance metrics using DPM and reference protocols.

In Chapter four, a channel allocation protocol with specific objectives, namely maximum throughput (MT), is constructed using the DPM mechanism. A queueing analytical framework that is able to obtain the performance metrics of each CR user independently is established based on DPM and the performance evaluations in a multiuser multi-channel CRN are carried out. The MT protocol is analyzed and compared with reference protocols also aimed at maximizing the throughput of the CR system with the numerical results.

In Chapter five, a study to investigate the optimal performance that a multiuser multi-channel CRN is able to achieve is carried out with the help of the DPM mechanism. The overall weighted throughput of the entire system under the delay constraints is selected as the optimization objective. The problem is formulated and solved under the framework of the DPM mechanism using a modified hill climbing algorithm proposed by the author.

Chapter six concludes the entire thesis and recommendations on future research are given.

CHAPTER 2 LITERATURE STUDY

2.1 CHAPTER OBJECTIVES

CR technology is aimed at achieving full use of limited radio resources using advanced technologies to meet the rapidly growing demands of wireless communication users. To implement CR technology in reality, it is crucial to solve the RA problem in CRN. Consequently, a considerable amount of research has been dedicated to addressing the RA problem in CRN. The design of the allocation protocol is an essential part of the research. However, various allocation protocols in related research works are either simply adapted from the existing protocols in the conventional communication system or modified to a limited extent. To clarify the research gap in the design of the RA protocol CR system, in this chapter, a comprehensive study of the prevalent research works on RA problems and design of the RA protocols is carried out. The bases of the RA design and modeling were introduced, followed by the method to evaluate the performance of the protocols. Existing protocols are categorized and introduced with their strengths and weaknesses highlighted. The study therefore demonstrates valuable directions for the CR RA protocol design and performance evaluation.

2.2 OVERVIEW OF COGNITVE RADIO TECHNOLOGY

CR technology has been studied for decades since the concept of CR was first introduced by Mitola in [27]. During the development and research of CR technology, there have been various selections of technologies in each link to implement CR technology in reality. It is redundant and unnecessary to cover every detail of CR technology in this thesis. In the remainder of this chapter, a brief introduction to the essential concepts of CR was given in this section. Then the details of the relevant technologies were discussed in the following sections. Readers or enthusiasts who needs more detailed information are hereby referred to the following references [28, 29, 29–31].

In the following part of this section, the CR functionality, spectrum access technique and CRN structure are introduced. Then the study subjects in this thesis and the reason why they have been chosen are discussed.

2.2.1 The cognition cycle

The CR functionality can essentially be represented by a cognition/cognitive cycle as shown in Figure 2.1, which consists of three main parts: observe, decide and act.

- **Observe:** The CR system is aware of the environment in which it is operating. With understanding of the communication demand of the CR users and understanding of the spectrum regulatory policies, the CR system carries out the measurement of the radio environment. Through the measurement, the following two types of information that represents the radio resources the CR system can utilize are obtained:
  - The maximum RF transmit power in the frequency band of interest that is regarded as having no harmful impact on the licensed user on the band at each transmitter of the CR system.
  - The estimation of the existence or absence of the licensed user in the frequency band of interest.
- **Decide:** With all the information the CR system gathers from the environment, the CR system makes decisions on how to utilize the available radio resources to fulfill the communication demand of the CR users. During the decision making, the following aspects are considered:
  - Estimation of channel-state information that indicates the channel capacity of the data transmission of the CR users.
  - Analysis of the potential harmful impact of the transmission of the CR user on the licensed users.
  - Determination of the duration of resource utilization scheme.

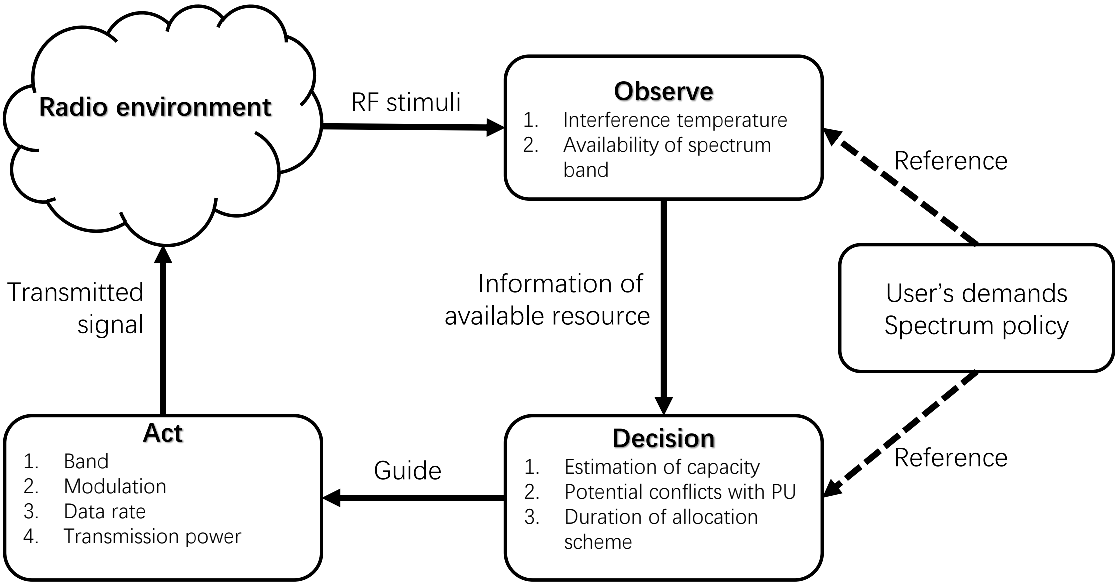


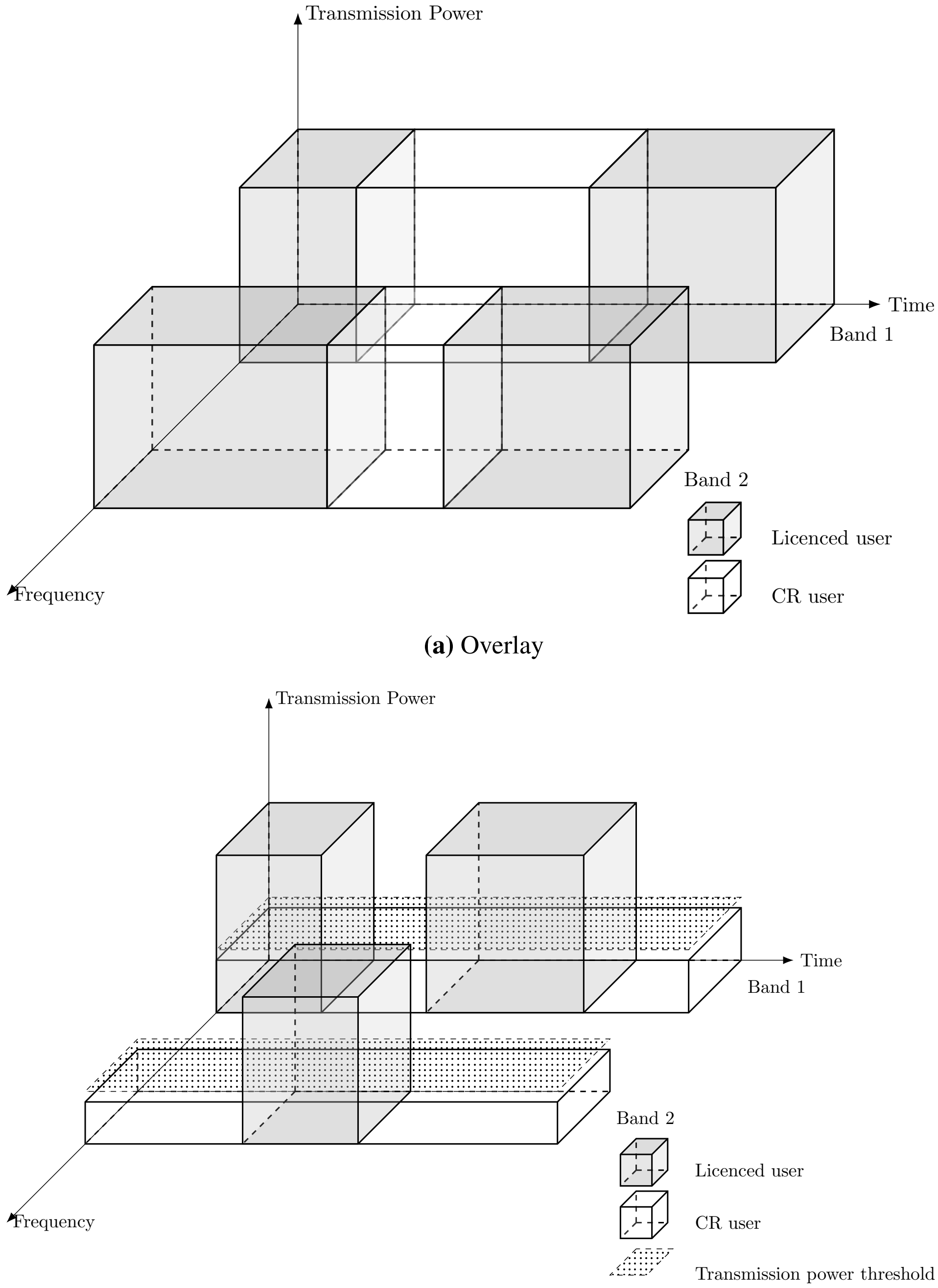
Figure 2.1. The cognitive cycle

- **Act:** The transmitter of CR users works according to the decision the CR system made. The following configurations of the transmitter can be adjusted to adapt to the dynamic decisions:
  - Operating frequency band
  - Modulation schemes and data rate
  - Transmission power.

2.2.2 Spectrum access technique

There are two main paradigms of spectrum sharing in the CR system, namely overlay and underlay [32–34]. An illustration of the two paradigms is shown in Figure 2.2 and the details are explained below.

- **Overlay spectrum sharing:** The CR user using the overlay spectrum sharing technique transmits through the spectrum band only when the spectrum band has not been used by its licensed users. As long as the CR system can detect licensed user's transmission accurately, interference with the licensed user is minimized.
- **Underlay spectrum sharing:** The CR user using the underlay spectrum sharing technique can transmit through the spectrum band at any time, as long as the transmission power is kept below



(a) Overlay

(b) Underlay

Figure 2.2. Spectrum access technique in cognitive radio.

a certain threshold so that the transmission of the CR user cannot make the transmission of the licensed user deteriorate.

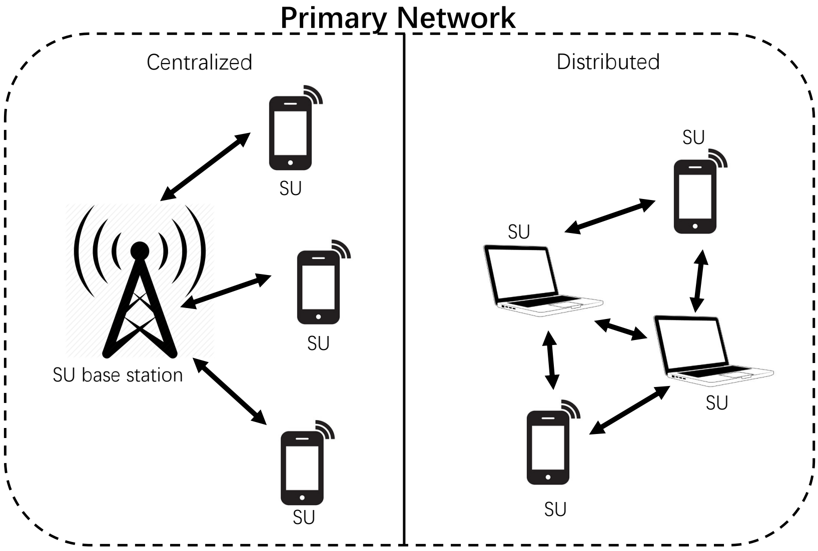


Figure 2.3. The centralized architecture and distributed architecture of CRN.

2.2.3 Cognitive radio network architecture

In the CR system, spectrum sensing and radio RA are key functionalities that must be implemented. Both of them require a dominant unit to collect acquired information and to make final decisions. There are two different architectures of the dominant unit and CR user devices in the CRN: centralized and distributed. Two examples of each architecture are shown in Figure 2.3. The details of both architectures are as follows [35–39]:

- **Centralized architecture:** In a centralized CRN architecture, the coverage of the CR service is divided into blocks in the form of cells or clusters. Inside each block, there is a center node that determines the behavior of the licensed users using the spectrum sensing information inside the block and allocates the available radio resource to the CR users inside the block. The center unit can be one of the CR user devices or a CR base station.
- **Distributed architecture:** In a distributed CRN architecture, the CR devices carry out the spectrum sensing and available resource analysis by themselves. Then they negotiate with each other using specific protocol on the utilization scheme of the radio resource.

2.2.4 Object of study in this thesis

The author chose the overlay CRN with centralized architecture as the study subject of this research. The considerations are as follows:

- In an overlay CRN, the conflicts of radio resources between the PU system and CR system and among the CR users are simple and clear to model. As discussed above, the result of the availability of certain radio resources to the CR system is binary (available or not) in an overlay CRN. Hence, the formulation of the available resources to the CR system is linear. It is convenient and efficient to analyze new ideas with this simple model first before in-depth study. Moreover, by reformulating the resource model, the discovery with an overlay model may be able to work as valuable references to the study of underlay CR system.
- The centralized architecture is applied with various considerations. In order to investigate analytical framework for complex networks, the basic and repetitive component of the network is usually studied first. Then, combining all components, the overall analysis of the whole network can be carried out. To model a basic component of a CRN, the centralized architecture is representative because it can describe the relationship between nodes inside the basic component comprehensively. The functionalities of the center node is virtual and can be implemented on every node in the network. Thus, the topology of centralized architecture can degenerate into links between two nodes by reducing the number of nodes to two. The links can be treated as the basic components of a distributed network. Similar to the situation above, the conclusions obtained by the centralized architecture can be adapted to suit the distributed architecture or provide references.

  Moreover, the virtual functionalities of the center nodes can serve as an interface to adapt advanced technologies that are applied in a distributed way, such as distributed spectrum sensing technology, by adjusting relative parameters at the center node. By introducing more complex models at the center node, this approach is able to achieve analysis of CR system with more precision.

Figure 2.4 shows an example of an overlay, centralized CRN working within the coverage of a PU network.

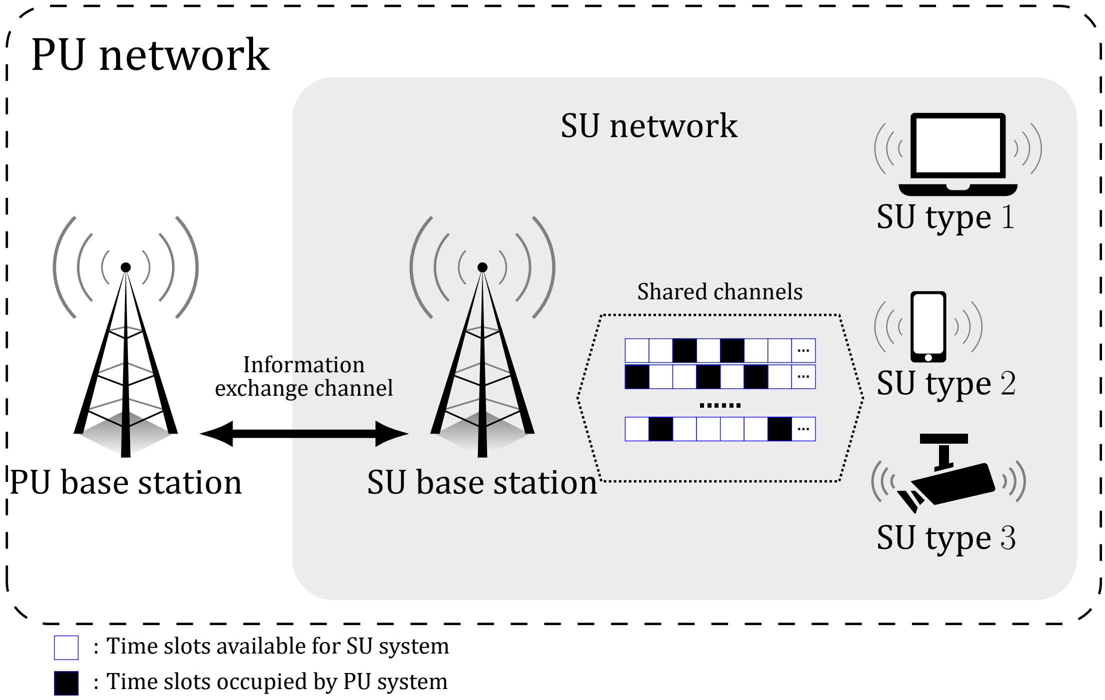


Figure 2.4. An example of the system model.

In the following section of this chapter, how to model such a system with every possible detail and technique is introduced.

2.3 METHODOLOGY TO MODEL COGNITIVE RADIO NETWORKS

2.3.1 Queueing theory and queueing model

Queueing theory is the mathematical study of queues and is generally considered a branch of operations research [40–43]. The origin of queueing theory is research on models to describe the Copenhagen telephone exchange by Agner Krarup Erlang. Queueing theory has been widely used in the research field of telecommunication [44–46].

The following features of queueing theory make it a suitable analytical tool for CR.

- Because the data transmission of the CR users is not guaranteed at any time, the data needs to be buffered when the radio resource is occupied by the licensed user. The buffer and the data restored in it can be considered a queue. Moreover, various kinds of QoS of CR user can be obtained through the analysis on this queue.

- Different from the traditional communication system, the CR system must be aware of some highly dynamic factors, such as the behavior of licensed users, the transmission demand of CR users and the condition of the CR user's data transmissions, to make good use of the limited radio resources. However the exact information of these factors are almost not obtainable because of the following conflicts:

  1. The rate of change and the delay of the information transmission.
  2. The rate of change and the limited device performance.

  This requires the CR system the ability to estimate these factors using priori knowledges on the evolution of the factors. Various probability models applied in the queueing theory, such as the Markov process, are suitable for this job.

2.3.2 Model of resources

The original concept of "resource" is something acting as source or supply, which can yield benefit if one can take advantage of it. There are two common features of resources in practice: limited availability and exclusivity. From the concept and features, an attempt was made to define "resource" in the CR system.

The most important exclusive factor in wireless communication is the frequency. It is the common rule for radio regulators worldwide to divide the whole spectrum by frequency into portions (frequency bands) and make them exclusive for the licensed users. To make full use of the assigned resource inside the frequency band, the radio resource is also divided into portions according to exclusive factors such as time, frequency, physical location, transmission power and coding. This technique is called multiplexing. The essence of CR technology is also a special form of sharing of the radio resource. The key exclusive factors are the same as above with one additional condition, which is avoiding harmful interference with licensed users.

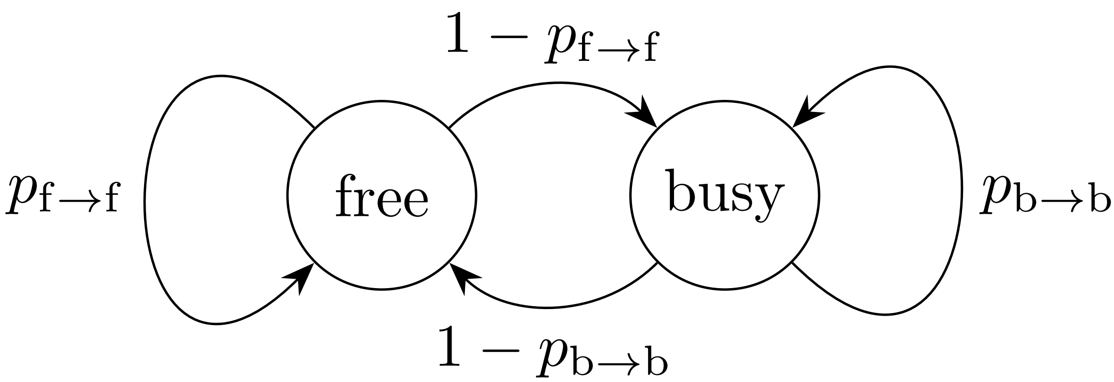


Figure 2.5. The Markov chain model of PU activity.

The resource is the surplus of the exclusive factors caused by the absence or the incapability of resource utilization of the licensed user. Because the behavior of the licensed user system is independent of the CR system, the resource that a CR system can utilize is time-varying and geography-varying. Because of the reasons mentioned in Section 2.3.1, the CR system is incapable of obtaining the real-time information of the available resource. Hence, the CR system needs a model that describes the dynamic feature of the resource to estimate the real situation.

The study subject in this thesis is the overlay CR system. The radio resource is formulated from two dimensions:

- **Channel:** All the exclusive factors that are irrelevant to time, such as the operating frequency, the modulation schemes and codings, are described abstractly as a concept "channel". The definition of channel is as follows: a channel is a combination of settings including the operating frequency, the modulation schemes and codings, etc. that can only be utilized to transmit data by one single user exclusively.
- **Time:** The channel is available to the licensed user all the time and is only available to CR users when it is not used by the licensed user.

The popular model to describe the behavior of the licensed user on one channel through time is a time-homogeneous discrete-time Markov chain of order 1 with two states [47]. One of the states represents that the channel is occupied by the licensed user, while the other represents that the channel is not occupied. To make the statements on the states unified and simple, in the rest of the thesis, the occupied state is named the "busy" state, and the unoccupied state is named the "free" state. The

Table 2.1. The parameters of PU activity model.

Category	Symbol	Description	Conversions
Original	$p_{\text{b}\to\text{b}}$	The probability that the current state is free given the previous state is free.	$t_{\text{PU}} = \dfrac{1}{1-p_{\text{b}\to\text{b}}}$
	$p_{\text{f}\to\text{f}}$	The probability that the current state is busy given the previous state is busy.	$\theta = \dfrac{1-p^{j}_{\text{f}\to\text{f}}}{2-p^{j}_{\text{b}\to\text{b}}-p^{j}_{\text{f}\to\text{f}}}$
Descriptive	t_{PU}	The average duration of the successive busy state.	
	θ	The percentage of the busy state. [48]	

model can be configured by a set of parameters to describe various kinds of licensed users' behavior patterns. The set of parameters can be divided into two categories: original and descriptive. The original parameters describe the transition probability between states and are shown in Figure 2.5. The descriptive parameters describe the average duration and the percentage of one state. The relationship between the two categories is shown in Table 2.1.

The discussion of the model of resources above is based on orthogonal multiple access. Non-orthogonal multiple access is a promising multiple access technique and has been proposed for 3GPP Long Term Evolution, the application of which will increase the complexity of the resource model. The research of the model of orthogonal multiple access can provide help and some basic knowledge when studying the non-orthogonal multiple access resource model.

2.3.3 Model of data transmission

To make full use of the limited radio resources, some advanced technologies are applied by the CR system. To improve throughput in wireless data communication systems, adaptive modulation and coding (AMC) have been advocated at the physical layer. At the data link layer, an automatic repeat request (ARQ) is usually used to counteract the channel fading, especially when forward error coding at the physical layer cannot meet the requirement. A cross-layer design that combines the physical layer and data link layer can optimize the performance, given constraints of both layers, such as prescribed delay or error rate. The details of the AMC and ARQ and the method to model the data transmission under the cross-layer design is introduced in the following part.

Table 2.2. Modulation and coding rates for IEEE 802.22

PHY mode	Modulation	Coding rate	Peak data rate in 6MHz (Mb/s)	Spectral efficiency BW = 6 MHz
1	BPSK	Uncoded	4.54	0.76
2	QPSK	1/2 and repeat: 3	1.51	0.25
3	QPSK	1/2	4.54	0.76
4	QPSK	2/3	6.05	1.01
5	QPSK	3/4	6.81	1.13
6	QPSK	5/6	7.56	1.26
7	16-QAM	1/2	9.08	1.51
8	16-QAM	2/3	12.10	2.02
9	16-QAM	3/4	13.61	2.27
10	16-QAM	5/6	15.13	2.52
11	64-QAM	1/2	13.61	2.27
12	64-QAM	2/3	18.15	3.03
13	64-QAM	3/4	20.42	3.40
14	64-QAM	5/6	22.69	3.78

Table 2.3. Modulation and coding rates for HPERLAN Type 2

Modulation	Coding rate	Nominal bit rate Mbit/s	Coded bits per sub-carrier	Coded bits per OFDM symbol	Data bits per OFDM symbol
BPSK	1/2	6	1	48	24
BPSK	3/4	9	1	48	36
QPSK	1/2	12	2	96	48
QPSK	3/4	18	2	96	72
16QAM	9/16	27	4	192	108
16QAM	3/4	36	4	192	144
64QAM	3/4	54	6	288	216

2.3.3.1 Adaptive modulation and coding

AMC is a technique that dynamically controls the modulation pattern and channel coding rate in order to achieve a better data rate [49]. The selection of the modulation pattern and channel coding rate is based on the information on the channel quality. If the channel quality is considered good, modulation of high order and high rate channel coding are employed and vice versa.

1 AMC STANDARDS In practice, each mode in the AMC consists of a specific modulation and coding pair. Table 2.2 shows the details of modulation and coding rates schemes in IEEE 802.22 [25,50]. Table 2.3 shows the PHY layer mode spec and parameters in HYPERLAN Type 2 [51–54].

2 CHANNEL MODEL Accurate channel models are essential for evaluating the performance of AMC. In research on modeling of wireless channels, frequency non-selective or flat-fading channels have been studied in depth and the model has been used for a variety of wireless communication applications. Understanding of flat-fading channels will be useful for future research on complicated frequency-selective fading channels. In this research, the wireless channels are all assumed to be flat-fading so that the mature method of modeling channels can be applied and accuracy of analysis can be guaranteed. The finite state Markov channel model is the most popular model to describe the flat-fading channels in wireless communication systems because of its versatility and capability of capturing the essence of time-varying fading channels with suitable choices of parameters. A brief history of the finite state Markov chain (FSMC) and its applications on fading channels, as well as related publications, is given in Table 2.4 [55].

Among the FSMC models, the first-order FSMC models are most widely used. A first-order Markovian assumption means that the current channel state is dependent on the previous state and independent of all other past states. The advantages of using first-order FSMC models are the computational simplicity and the ease of determining the model parameters. Although there is no consensus about the situation in which first-order FSMC models are accurately applicable to fading channels, it is a common view that first-order FSMC models are acceptable for slow fading channels.

The received signal-to-noise ratio (SNR) is a common parameter of a fading channel that is used to map the channel conditions to FSMC states. The FSMC modeling of the received SNR works as follows: A channel is partitioned into K non-overlapping regions according to the received SNR γ such as $[0,\Gamma_0)$, $[\Gamma_0,\Gamma_1)$, $[\Gamma_{K-2},\infty)$. There are three widely used statistical models for fading SNR in the FSMC modeling: Rayleigh [77–79], Rician [80–82] and Nakagami m [83–85]. The probability density functions (PDF) of the SNR of the three models are listed in Table 2.5.

There are several considerations when selecting the boundary of the SNR regions. First, the SNR range of each region should be wide enough to make sure that most of the time, the SNR variation within the duration of a received packet stays in only one region and the SNR variation within the duration of the following packet stays in the current region or one of the two adjacent regions. On the other hand, the SNR range should not be made too large because if the SNR range increases, different packets falling in one region may have different bit error rate (BER), which make the analysis based on the FSMC model and BER inaccurate.

Table 2.4. A brief history of FSMC and its applications for fading channels [55]

Year	Contributions	Research papers
1957-1958	Coding theorems and proof of channel capacity for finite-state channels	[56], [57]
1960-1963	Capacity and bit error rate of burst-noise telephone circuits using a two-state Markov model	[58], [59]
1963-1968	extension of Gilbert-Elliott channel models	[60], [61]
1968	Capacity and coding theorems for indecomposable FSMC	[62]
1969-1989	Applications of FSMC models to wire-line circuits or wireless links between fixed stations	[63]
1989	Capacity and decoding for the Gilbert-Elliott channel	[64]
1991-1995	Early FSMC modeling of UHF channels and FSMC-based adaptive coding	[65], [66]
1995	FSMC modeling of the received SNR based on the Clarke's model	[67]
1997-2006	Modeling error bursts in fading channels using FSMC and system performance evaluation	[68], [69], [70]
1996-2006	Decoding and estimation in FSMCs or in fading channels that are modeled as FSMCs	[71], [72]
1991-2006	Adaptive transmission in fading channels using FSMC models	[66], [73], [74]
1996-2006	Accuracy and applicability of first-order and higher-order FSMC models for fading channels	[68], [75], [76]
1996-2006	Information theory of FSMC models (with or without feedback)	[71]

There are two partition methods for the SNR region.

1. The equal-duration method: Select a group of SNR boundaries $\Gamma_0, \ldots, \Gamma_K$ that creates the expectation of the duration that the SNR falls in each region being identical.
2. The equal-probability method: Select a group of SNR boundaries $\Gamma_0, \ldots, \Gamma_K$ that creates the probability that the SNRs fall into each region being identical.

The equal-duration method is more popular because the expectation of the duration can be set to be some multiple of the packet length so that the following analysis based on the FSMC model can be carried out with ease.

Let π_k denote the steady-state probability of k-th region.

$$\pi_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p_\Gamma(\gamma) d\gamma. \tag{2.1}$$

Table 2.5. PDF of statistical models of fading channel

Name	PDF	Parameters
Rayleigh	$p_\Gamma(\gamma) = \frac{1}{\rho}\exp\left(-\frac{\gamma}{\rho}\right), \gamma \geq 0$	γ: SNR ρ: average SNR
Rician	$p_A(a) = \frac{2aK}{a_s^2}\exp\left(-\frac{K(a^2+a_s^2)}{a_s^2}\right) I_0(\frac{2Ka}{a_s^2})$	a: received signal amplitude a_s: average amplitude K: Rice factor $I_0(\cdot)$: the zeroth-order modified Bessel function of the first kind
Nakagami m	$p_\Gamma(\gamma) = (\frac{m}{\rho})^m \frac{\gamma^{m-1}}{\Gamma(m)} \exp(-\frac{m\gamma}{\rho})$	γ : SNR ρ : average SNR m: shape parameter $\Gamma(\cdot)$: the gamma function

The level crossing rate of Γ for the SNR process.

$$N(\Gamma) = \sqrt{\frac{2\pi\Gamma}{\bar{\gamma}}} f_{\mathrm{D}} \exp(-\frac{\Gamma}{\bar{\gamma}}), \tag{2.2}$$

where f_{d} is the maximum Doppler frequency shift caused by the motion of the transceiver. Then the average duration of the SNR region $[\Gamma_k, \Gamma_{k+1})$ can be obtained as

$$\bar{\tau}_k = \frac{\pi_k}{N(\Gamma_k) + N(\Gamma_{k+1})}. \tag{2.3}$$

Let T_P denote the duration of a packet. The transition probability can be approximated as

$$P_{k,k+1} \approx \frac{N(\Gamma_{k+1})T_{\mathrm{P}}}{\pi_k}, \tag{2.4}$$

$$P_{k,k-1} \approx \frac{N(\Gamma_k)T_{\mathrm{P}}}{\pi_k}, \tag{2.5}$$

where $P_{j,k}$ is the transition probability from the j-th region to the k-th region.

The FSMC model is able to represent some actual channels in practice. For example, from data obtained during an indoor broad-band measurement campaign [86], in microcellular/picocellular environments channel behavior can be represented by the Ricean model. Despite this fact, the model of fading channel mentioned above may be not able to represent practical fading channels with both time-varying and frequency-selective characteristics. It is a direction for future research to extend the FSMC model to include frequency-selective fading channels of various types.

2.3.3.2 Automatic repeat request

ARQ is an error-control method at the data link layer for data transmission over an unreliable channel [87–90]. It uses messages sent by the receiver indicating if the data packets have been received correctly or not. This message is usually called acknowledgments (ACK). The commonly used ARQ techniques are as follows:

- Stop-and-wait ARQ: When stop-and-wait ARQ is applied, a sender transmits one single packet and then waits for the ACK from the receiver that indicates the status of the packet. If the ACK indicates that the original packets was lost or corrupted, the sender will repeat the transmission. Otherwise, the sender continues transmission of the next packet.
- Go-back-N ARQ: When go-back-N ARQ is applied, each packet is labeled by a unique sequence number. The sender continues to send N packets even without an ACK from the receiver. The receiver keeps track of the sequence number of the next packet it expects to receive, and sends the ACK indicating the number of the last correctly received packet. The receiver only accepts the packet with the correct sequence number that is after the sequence number of the last correctly received packet, and discards all packets with other sequence numbers. Once the sender has sent all N packets, it will go back to send the packet with the sequence number of the last ACK it received. The whole process is repeated.
- Selective repeat ARQ: When selective repeat ARQ is applied, each packet is also labeled with a unique sequence number. The receiver keeps track of the sequence number of the earliest packet it has not received, and sends ACK with that number. The transmitter continues to send N packets, no matter what ACKs it receives. The receiver continues to accept the subsequent packets, replying each time with an ACK indicating the sequence number of the earliest missing packet. Once the sender has sent all N packets, it retransmits the packet with the number indicated by the ACKs, and then continues where it left off.

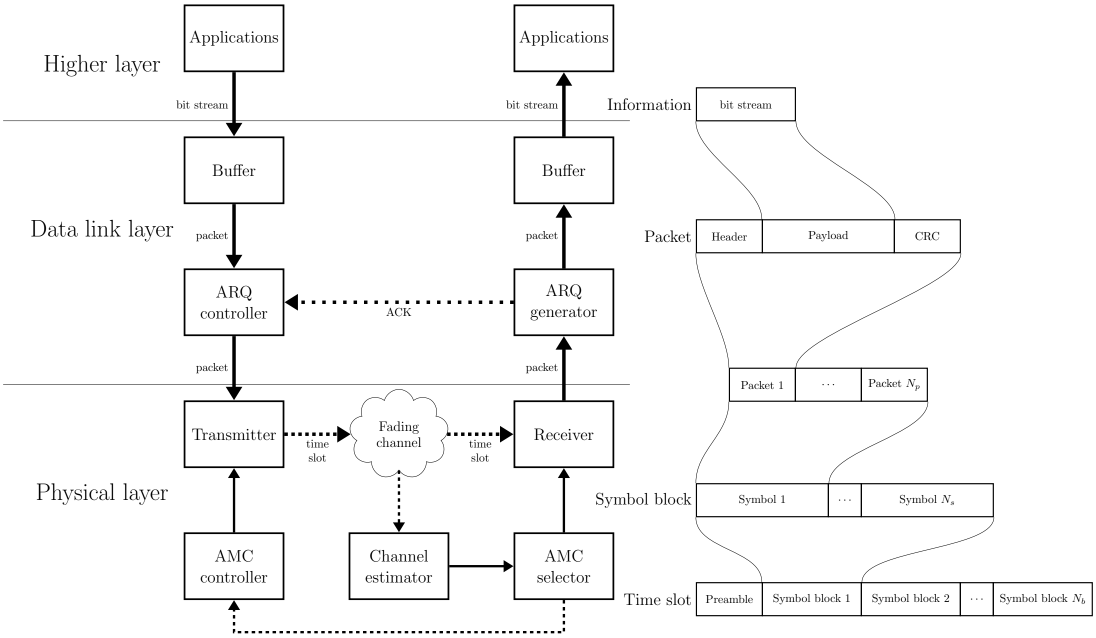


Figure 2.6. The cross-layer design.

2.3.3.3 Cross-layer design

When the performance related to the queueing effects needs to be analyzed and optimized, such as the throughput, delay and error rates, the AMC on the physical layer and ARQ on the data link layer should be considered and designed together. This approach is called "cross-layer design" [89, 91–94].

In a CR system using cross-layer design as shown in Figure 2.6, the data from higher layers that are required to be transmitted are in the form of bit streams. The bit streams are mapped to packets. The number of bits that each packet contains is fixed and denoted by N_{bit}. A typical structure of a packets includes the packet header, payload, and cyclic redundancy check bits. The channel state information (CSI) estimation is carried out by the CR system and modulation and coding pair is chosen according to the CSI estimation. The bit stream in each packet is modulated in the form of symbols. The number of bits contained in each symbol is determined by the modulation and coding mode chosen and denoted by R_{s} (bits/symbol). Thus, the packets are mapped to a block of $N_{\text{bit}}/R_{\text{s}}$ symbols.

At the physical layer, transmissions are operated with the structure of the time slot (or frame). The duration of the time slot is constant and denoted by T_{t}. A fixed number of symbols are transmitted in one time slot.

At the data link layer, ARQ is applied to guide the retransmission of the packets that are not correctly received. The data storage ability of the system is represented in the form of a buffer at the data link layer. The data awaiting transmission are treated in the form of a packet and stored in the buffer at the transceiver. Here, considering that in practice the buffer size and the delay caused by the ARQ retransmission all have limitations, there should be an upper bound of the ARQ retransmissions, which is denoted by N_{ARQmax}.

Taking the MPEG-4 audio/video transmissions as an example, the quality of service requires delays no more than 150 to 400 ms [95]. Assuming that the average round trip delay between the sender and the receiver is 100 ms, the number of ARQ retransmissions should not exceed four.

If one packet is not received correctly after N_{ARQmax} retransmissions, the packet will not be transmitted again.

2.3.4 Evaluation of the performance

The evaluation of performance metrics can be carried out with the queueing analysis framework. The performance metrics that attract the researchers' interest are as follows:

- Throughput
- Delay
- Queue length
- Packet transmission failure

  - Packet loss
  - Packet rejection
  - Packet collision

2.3.4.1 Throughput

The common definition of throughput is that throughput is the rate of successful message delivery through a communication channel or from a communication device. Throughput is one of the most important performance metrics from both the user's and the service provider's perspective. The throughput represents the number of messages that can be sent or received by the service of the communication system from the user's perspective. On the other hand, throughput represents the extent of message delivery in total that can be provided to potential customers by the communication system from the service provider's perspective.

In the queueing analytical framework, the message is usually measured in the form of packet at the data link layer. Thus, the throughput is represented by the rate of successful packet delivery.

2.3.4.2 Delay

Delay, sometimes known as latency, is a concept in a packet-based communication network that represents the time interval between the occurrence of a message transmission demand and the completion of the transmission.

The delay of a tagged packet comprises two parts:

1. **Waiting time**: When the packet with the message required to be transmitted is generated and ready to be transmitted, there might be some other packets that have been generated ahead of it waiting in the buffer of the sender. If the tagged packet has the same or a lower priority as the packets ahead (such as the first in first out rule), it will not be transmitted until all the packets ahead have been transmitted. Thus, the waiting time is the time interval between the generation of the tagged packet and the start of its transmission.
2. **Transmission time**: If the transmission is done in single hop, the transmission time is the time of the propagation of the signal carrying the packet from the sender to the receiver. If the transmission is done in multiple hops, besides the propagation time of the multiple hops, the time of the routing of the packet at the relay nodes is included in the transmission time.

In CR system, the transmission of the SU is not guaranteed. As a result, the waiting time part of the delay is the focus of most studies of CR system.

2.3.4.3 Queue length

The transmission of SUs in the CR system is not guaranteed because the utilization of the radio resource is limited to avoid harmful interference with the PU system. Thus, a buffer that restores the packets awaiting transmission while the radio resources are not available to the SU system is requisite at the sender. Among the studies on the CR system using the queueing analysis framework, the buffer is a crucial subject and most of the analyses of the system are based on the buffer. The queue length is the essential attribute of the buffer, which means the number of packets waiting in the buffer. The investigation of the number of packets in the buffer is able to reveal various performance metrics, such as the throughput, the delay and the packet transmission failures. The relationship obtained can help design the size of the buffer.

2.3.4.4 Packet transmission failure

As introduced in Section 2.2, the data transmission in the CR system requires multiple procedures, such as spectrum sensing and RA. The effects of those procedures on data transmission is more complex than the traditional communication system and will cause various types of packet transmission failures. The packet transmission failures that attract the interest of researchers are as follows.

- **Packet loss**: The packet loss in the CR study is caused by the ARQ retransmission limit mentioned in Section 2.3.3.2. The packet loss is usually measured in term of the packet loss rate, which is generally represented by the average number of packet losses in one time unit (e.g. one time slot).
- **Packet rejection**: Packet rejection happens when a packet transmission demand is generated and meanwhile the buffer is full. The system fails to store the packet in the buffer and rejects the transmission demand of the packet. The packet rejection is usually measured in terms of the packet rejection rate, which is generally represented by the average number of packet rejections in one time unit (e.g. one time slot).

- **Packet collision**: When the CR system fails to sense the PU activity correctly and transmits in an illegal manner, such as using excessively high transmission power or transmitting through a channel already occupied by the PU, the packet transmitted by the CR system causes collision with the PU's transmission. Since the collision packet might be successfully received by the SU receiver (however it will cause interference with the PU system), the packet collision is a special type of transmission failures. The packet collision is usually measured in terms of the packet collision rate, which is generally represented by the average number of packet collisions in one time unit (e.g. one time slot).

2.4 RESOURCE ALLOCATION PROTOCOLS

With the development of CR technology, the CR system will be capable of utilizing radio resources while avoiding harmful interference with PU systems. Optimize the RA in CRN is an important issue [96–98]. It is reasonable to assume that a CR system can utilize radio resources on a larger scale than a traditional communication system. Generally, the radio resources that a CR system can utilize are divided into multiple non-overlap parts, namely channels (in Section 2.3.2). The way in which the CR system utilize the channels is a crucial factor in the system performance. The RA protocol is a set of criteria that the CR system uses to determine how to assign the available channels to the CR users.

The allocation protocols based on the framework mentioned in the literature can be divided into two categories: channel independent protocols and channel dependent protocols. These two types of protocols are discussed in the subsequent sections.

2.4.1 Channel independent protocols

The channel independent protocols are the allocation protocols that do not consider the channel conditions, which may affect the transmission rate. Most channel independent protocols are based on the following principle: Each user in the system has the same opportunity to access the available channels, regardless of any other circumstances. When the channel independent protocol is applied by a CR system, the channel allocator of the CR system simply needs to consider the availability of the PU channels and the transmission demands of the CR users. Thus the application of the channel independent protocol does not require the channel allocation to have much computational ability compared to other protocols because the channel conditions are ignored during the allocation process. On the other hand, the randomness of the allocation results yield worse performance than the other protocols.

The application of the channel independent protocol is generally under the following conditions.

- The computational ability of the CR system is limited.
- The requirement of the performance of the CR system is not high.
- The radio resource is sufficient for the CR users' demand.

There are two methods to implement the random allocation protocol: random allocation [47] and round robin [99–101].

- **Random allocation**: There is a random number generator at the resource allocator. Each time the available channels are assigned, the allocation is determined by the random number generated in a special way so that each SU shares the identical opportunity to assess all PU channels.
- **Round robin**: RR is originally one of the algorithms employed by process and network schedulers in computing to assign computational resources in terms of time slices to each process in circular order. All processes are handled without priority and the access of each process to the resources is guaranteed, so that no process will suffer task congestion caused by inaccessibility of the resources. RR allocation protocol is able to achieve maximum-minimum fairness if the processes are the same size.

  RR can be adapted to the CR system if the computational resources are replaced by radio resources and the processes are replaced by the packet transmission demand. Nevertheless, the results that RR in the CR system could achieve are more complex, because the transmission capacity that the radio resources can provide to each other may vary.

2.4.2 Channel dependent protocols

As mentioned above, the channel independent protocol is unable to achieve optimal performance and high radio resource utilization efficiency because the channel conditions are ignored during the allocation. Therefore, in order to achieve higher performance and radio resource utilization efficiency, the consideration of channel conditions has to be contained in the mechanism of the allocation protocols. This type of protocol that allocates the channels taking into consideration the information of the channel conditions, which is also mentioned as the CSI in the related literature, is classified as channel dependent protocol.

Using the CSI, various objectives can be aimed at by the channel dependent protocols, such as throughput, delay, fairness and failure rate. Among these objectives, throughput maximization has drawn a lot of attention and is also prevalent in the study of CR RA. Most of the protocols aimed at maximizing throughput in a CR system are adapted from the opportunistic scheduling protocol [48, 102–106].

The concept of opportunistic scheduling was introduced by Knopp and Humblet in [107]. Their study showed that capacity can be significantly improved using multi-user diversity in scheduling process. The original opportunistic approach is that the allocator always assigns the resources to the users who have the best conditions to use the resources and therefore have the best potential throughput. This approach is usually referred to as the max rate. Furthermore, the criterion used to compare the channel conditions of different users can be modified to adapt some complicated situations when the CR system intends to pursue other objectives besides maximizing throughput, such as the fairness of transmission chance. The method to achieve other objectives is generally based on the method to maximize throughput. For instance, a max weight approach [108, 109] uses a weight that is relevant to the number of packets in the buffer. The allocator assigns the resource to maximize the product of the weight and the throughput. Thus, blocking of the packets can be avoided.

The channel dependent protocols are usually applied when the resources are insufficient for all users' demand. This scenario is common in the CR system. As a result, the channel dependent protocols, especially opportunistic scheduling, are generally applied in the channel allocation design of the CR system.

2.4.3 Weakness of the existing protocols

The protocols adapted from the conventional communication system have been successfully implemented in the CR system through proper modifications. However, there is a potential broad development space for the allocation protocol to be improved and perfected. According to the author's observations, improvements based on the following features that the existing protocols lack warrant further study.

- Applicability to varying objectives. The application environment of CR technology is volatile and complex and diversities are caused by both the characteristics of the PU systems that own the spectrum the CR system is using and the various possibilities of the demands of potential CR users. This requires that the set of allocation protocols from which the allocator can choose is capable of fitting diversified and varying objectives of the RA. However, the protocols mentioned above are either of no relevance to the objectives, such as the random allocation and RR, or merely capable of fitting restricted and fixed objectives, such as opportunistic scheduling.
- Adaptability and flexibility of configuration. Besides the diversity of the feasible set of possible objectives of allocation, the time-varying feature is another obstacle to design a channel allocation protocol. Both the changing of the objectives caused by the involvement of new users and/or the exit of existing users and the changing of the available resources caused by the PU activities require the protocols to be adaptable to these situations and the switches among different protocols to be simple and effective.

2.5 OPEN PROBLEMS IN RESOURCE ALLOCAITON PROTOCOL DESIGN

In summary of the literature review in this chapter, indications of possible solution to channel allocation protocol design problems are given below:

- Identify a mechanism to design channel allocation protocols for CRN that are applicable to varying objectives and offer flexibility of configuration.
- Establish a framework that is able to model various types of channel allocation protocols and take them as an input to analyze system performance. Furthermore, using such a framework, find a convenient method to compare the performance of different protocols.

- With the help of the channel allocation design mechanism and the framework for performance evaluation, design a procedure to configure/create optimal allocation protocols given various objectives.

2.6 CONCLUSIONS

In this chapter, the analytical framework to investigate the design and evaluation of the CR channel allocation protocol with the purpose of addressing the RA problem in CRN was reviewed. The bases and assumptions that validate the framework were organized and introduced. The features that enable channel allocation protocols to be applied in CRN were investigated and the research gaps in designing and modeling channel allocation protocol and performance evaluation were identified. The conclusions and ideas in this chapter helped to identify further research done by the author to bridge the gaps, which are introduced in the remaining part of the thesis.

CHAPTER 3 MECHANISM FOR COGNITIVE RADIO CHANNEL ALLOCATION: DISTRIBUTION PROBABILITY MATRIX

3.1 CHAPTER OVERVIEW

The reason why CR technology is a promising wireless communication paradigm is its capability to mitigate the radio resource scarcity challenge and underutilization problem. To achieve these promises, an appropriate framework to allocate the limited resources to numerous CR users prudently is indispensable. The framework has two essential tasks: first, a proper model to describe the resource that the CR system is allowed to utilize must be established. Second, based on the model of the resources, the CR system has to implement an effective mechanism to create RA protocols that the allocator can employ to fulfill diverse requirements. The protocol creation mechanism must be aimed at the characteristics of the CR system to make the protocols convenient to operate and effective to ensure the fulfillment of preset objectives.

In this chapter, a protocol creation mechanism developed by the author, which is aimed at achieving the objectives above is introduced. The protocols generated by the mechanism are flexible, configurable and fit in well with the queueing analytical models. The validity and versatility of the mechanism are demonstrated by examples of applications in CRN and correspond with numerical results in the next part of this chapter.

3.2 INTRODUCTION TO THE DISTRIBUTION PROBABILITY MATRIX

3.2.1 Concept and basis of distribution probability matrix

Both the internal and external working environment of the CR system are highly dynamic. A channel allocation protocol that is aimed at achieving optimal RA needs to adapt its allocation strategy to the following dynamic conditions:

1. The availability of radio resources, which is affected by the PU activities.
2. Transmission conditions, which affect the transmission rate of CR users.
3. Varying demands caused by the diversity of users that the CR system serves.
4. The overall objective of the allocation protocol that balances the limited resources (obtained from point 1 and 2 above) and the various demands of users (from point 3).

To meet the requirements above, the author developed a mechanism to describe the allocation strategy: DPM [110]. This mechanism comprises two key elements: the allocation indicator and the effective duration, which are introduced below.

- **The allocation indicator**. In the mechanism, a parameter is created to indicate the possibility of the allocation of the resources during a certain period. Here, it is assumed that the utilization of the resource is exclusive for a single user. Let $\boldsymbol{R} = \{r_1, r_2, \dots, r_J\}$ denote the total J minimal non-overlapping units of resources that a CR system can utilize and $\boldsymbol{U} = \{u_1, u_2, \dots, u_M\}$ denote the total M users that the system serves. Hence, the implementation of such a parameter is as follows: the possible allocation of the j-th unit is measured by a group of probability variables $\boldsymbol{D}^j = \{d_0^j, d_1^j, \dots, d_M^j\}$, where d_i^j represents the probability that the j-th unit is assigned to the i-th user and d_0^j represents the probability that the j-th unit is not assigned to any user. Thus, the possible allocation of all the resources can be presented by a matrix, namely the DPM, which is defined below.

$$\boldsymbol{P}_{\text{DPM}} = \begin{pmatrix} \boldsymbol{D}^1 \\ \vdots \\ \boldsymbol{D}^J \end{pmatrix} = \begin{pmatrix} d_0^1 & d_1^1 & \cdots & d_M^1 \\ \vdots & \vdots & \ddots & \vdots \\ d_0^J & d_1^J & \cdots & d_M^J \end{pmatrix}. \tag{3.1}$$

According to the assumptions, in $\boldsymbol{P}_{\mathrm{DPM}}$: $d_i^j \in [0,1]$ for any $i \in [0,M]$, $j \in [1,J]$ and $\sum_{i=0}^{M} d_i^j = 1$ for any $j \in [1,J]$. There is also a simplified version of DPM that only considers the probability that one channel is assigned to the user, which is defined below.

$$\boldsymbol{P}_{\mathrm{DPM}} = \begin{pmatrix} \boldsymbol{D}^1 \\ \vdots \\ \boldsymbol{D}^J \end{pmatrix} = \begin{pmatrix} d_1^1 & \cdots & d_M^1 \\ \vdots & \ddots & \vdots \\ d_1^J & \cdots & d_M^J \end{pmatrix}. \tag{3.2}$$

- **The effective duration**. The allocation indicator, $\boldsymbol{P}_{\mathrm{DPM}}$, provides great flexibility for the resource allocator to formulate the allocation strategy. The next essential task is to determine the effective duration, which means the time when a particular $\boldsymbol{P}_{\mathrm{DPM}}$ generated by certain allocation strategy is valid. The choice of the effective duration is related to the computing capability of the allocator and the demand for flexibility. The shorter effective duration provides more flexibility in adapting different allocation strategies to suit the time-varying environment changes. However, a shorter effective duration requires the allocator to alter the $\boldsymbol{P}_{\mathrm{DPM}}$ more frequently, thus limiting the time for calculation that transfers the allocation strategy to the $\boldsymbol{P}_{\mathrm{DPM}}$. The trade-offs between the demand for flexibility and computing capability must be made by determining the effective duration of each $\boldsymbol{P}_{\mathrm{DPM}}$.

3.2.2 Application of the DPM

According to the characteristics, the DPM is able to serve the following purpose in the CR study.

- **Channel allocation protocol**. When a CR system has multiple channels to assign among the users and diverse objectives of channel allocation are required, the DPM is an option that can provide flexibility to meet the requirements. By setting up the allocation indicator $\boldsymbol{P}_{\mathrm{DPM}}$ and the effective duration, the channel allocator can allocate multiple channels to multiple users with the guidance of the DPM. The freedom of the configuration of the $\boldsymbol{P}_{\mathrm{DPM}}$ and the selection of an effective duration ensures the ability of the DPM to behave versatilely as a channel allocation protocol.
- **Protocol analytical tool**. In the study of the design of channel allocation protocol for multi-channel multi-user CRN, miscellaneous protocols are based on various ideas and implemented in diverse ways. Hence, it is difficult to compare the performance of these protocols without a unified analytical framework. The flexibility of configuration enables the DPM to describe

and implement the ideas behind other protocols. Thus, the protocols can be simulated by the DPM using corresponding setup and the performance comparison can be carried out in a fair and reasonable manner.

3.3 CHANNEL ALLOCATION PROTOCOL BASED ON DPM

One important application of the DPM is serving as a channel allocation protocol for the CRN. In the following part of this chapter, a typical application is given and the novelty of the mechanism is revealed by the numerical results and comparison with other channel allocation protocols.

3.3.1 System model and assumptions

In this section, an infrastructure-based CRN with multiple SUs and multiple channels that is shown in Figure 3.1 is considered. A CR system works under the coverage of a PU system and serves the SU. The CR system is called an SU system in the rest of the thesis. Both PU and SU systems are composed of one base station and multiple mobile stations. There is a reliable channel between the PU and SU base stations to exchange all the necessary CSI, including the channel condition, spectrum sensing results and synchronization information [111].

All the data transmitted in this system are assumed to be in the form of packets, which are organized and buffered at the data link layer. The packets are transmitted within a physical layer frame. It is assumed that all the physical layer frames in PU and SU systems are synchronized and of identical time duration T_{f}, which is defined as one time slot. The spectrum band authorized to the PU system is divided into J channels. The SU system shares the channels in an overlay mode introduced in Section 2.2.1: when the channel is not occupied by the PU system, the SU system can use it to transmit its packets. The data packets of the SU system are stored in a finite buffer at each SU mobile station and only the uplink data transmission from SU mobile station to base station is considered. To adapt this research for downlink transmission, the packet queues for SU should be located at the SU base station and each SU is assigned a separate buffer to restore its packets.

The assumptions above are based on the condition that the PU system is friendly and cooperative with the CR system. Thus, the cost to obtain important control information for the CR system, such

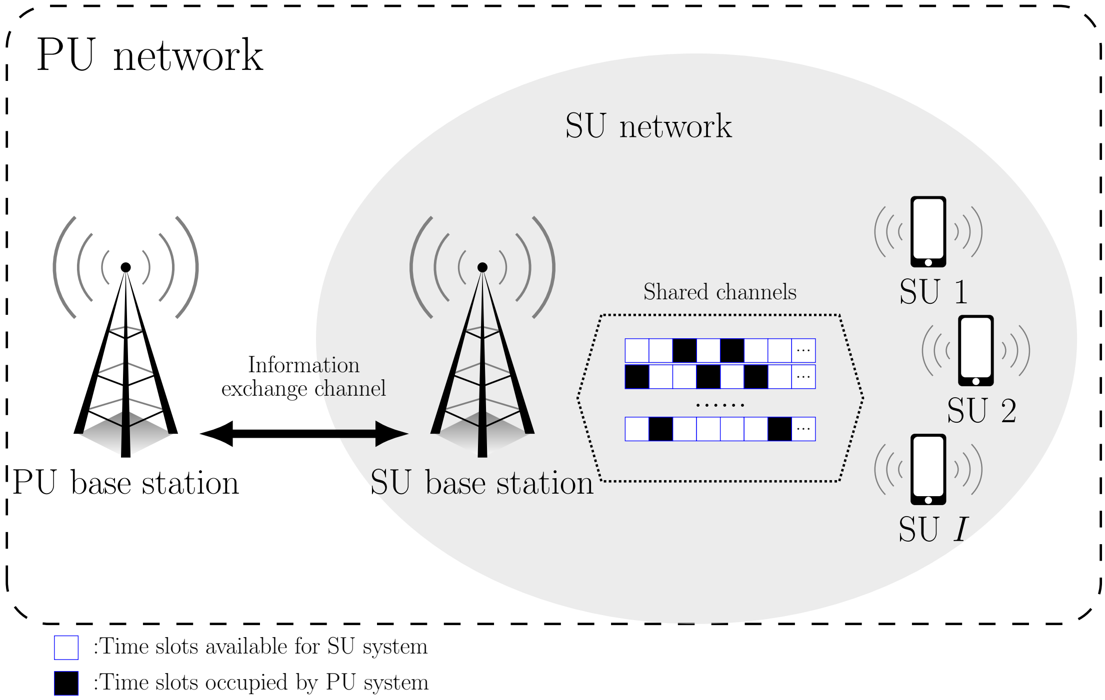


Figure 3.1. System model.

as the information for synchronization with PU system, can be ignored. The rest of this thesis is all based on this assumption. Non-cooperative property of PU system will introduce extra performance deterioration, which could be investigated in the future research.

3.3.2 Resource model

As introduced above, the radio resources that the CR system can utilize are divided into J non-overlapping channels and the availability of each channel depends on the PU transmission through it. The minimal time unit of the utilization of a channel is one time slot T_{f}. Thus the model of the radio resource of the CR system can be represented by the model of the PU transmission activities on each channel in each time slot.

The PU transmission activities on a channel are modeled as a time-homogeneous first-order Markov process with two states. Let $O^j(t) \in \{0,1\}$ represents the PU occupancy state of the j-th channel at the t-th time slot. When a PU is transmitting, the state is called "busy" and $O^j(t) = 0$, otherwise, the

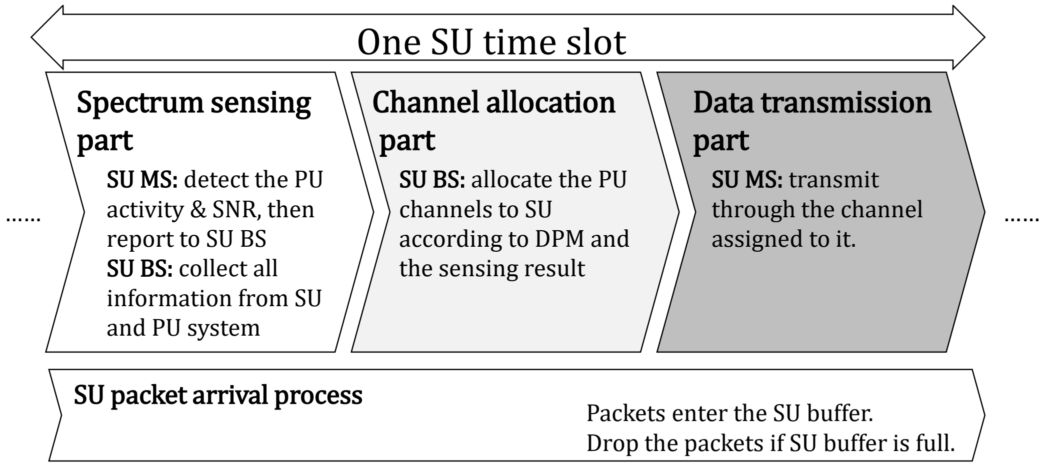


Figure 3.2. The time slot structure of the SU system.

state is called "free" and $O^j(t) = 1$. The process can be described by a transition matrix:

$$\boldsymbol{P}_{\mathrm{PU}}^{j} = \begin{pmatrix} p_{\mathrm{b}\to\mathrm{b}}^{j} & 1 - p_{\mathrm{b}\to\mathrm{b}}^{j} \\ 1 - p_{\mathrm{f}\to\mathrm{f}}^{j} & p_{\mathrm{f}\to\mathrm{f}}^{j} \end{pmatrix}, \tag{3.3}$$

where $p_{\mathrm{b}\to\mathrm{b}}^{j} = \Pr(O^j(t) = 0 | O^j(t-1) = 0)$ denotes the probability that the j-th channel is occupied by PU (the "busy" state), given that the channel is occupied by the PU (the "busy" state) at the previous time slot, and $p_{\mathrm{f}\to\mathrm{f}}^{j} = \Pr(O^j(t) = 1 | O^j(t-1) = 1)$ denotes the probability that the j-th channel state becomes "free" from a "free" state. These probabilities can be obtained by either measurement or estimation of the PU transmission behavior.

3.3.3 Secondary user system transmission model

Each SU time slot consists of three successive parts: spectrum sensing, channel allocation and data transmission as shown in Figure 3.2.

3.3.3.1 Spectrum sensing

At the beginning of each time slot, the SU system senses the existence of the PU transmission on every channel. Let $\hat{O}^j(t) \in \{0, 1\}$ denote the sensing result. $\hat{O}^j(t) = 1$ means the SU system estimates that the j-th channel is not occupied by the PU system at the t-th time slot, and the SU system will allocate this channel to one of the SU mobile stations and transmit data between the base station and the mobile station. $\hat{O}^j(t) = 0$ means the SU system estimates that the j-th channel is occupied by the PU system at the t-th time slot and the SU system will not use the channel during this time slot. In this model, imperfect spectrum sensing is considered. There are two kinds of sensing errors: miss detection and false alarm. $P_{\text{md}} = \Pr(\hat{O}^j(t) = 1|O^j(t) = 0)$ represents the probability of miss detection, which means the SU system has the sensing result of "free" state while the PU is actually transmitting. $P_{\text{fa}} = \Pr(\hat{O}^j(t) = 0|O^j(t) = 1)$ represents the probability of a false alarm, which means the SU system gives the sensing result of "busy" state while the PU is not transmitting. If miss detection happens, the packets that the SU has transmitted on the channel at that time slot will interfere the PU data transmission and are defined as "collision" packets.

3.3.3.2 Channel allocation

After the spectrum sensing part, the SU system allocates all the channels that are sensed "free" to the SU mobile stations. Here, the DPM is applied. The protocol works as follows: First, an N-by-M DPM is set up as indicated below.

$$\boldsymbol{P}_{\text{DPM}} = \begin{pmatrix} d_1^1 & \cdots & d_1^M \\ \vdots & \ddots & \vdots \\ d_N^1 & \cdots & d_N^M \end{pmatrix}, \tag{3.4}$$

where d_i^j represents the probability that the j-th channel is allocated to the i-th SU at the beginning of a time slot if the channel is sensed "free". In this framework, it is assumed that one SU can transmit on multiple channels, but one PU channel can be accessed by only one SU. So in P: $\sum_{j=1}^{M} d_i^j$ for any i.

At this stage of the study, the author's focus is investigating the potential performance improvement of the DPM itself as the channel allocation indicator. In order to eliminate the influences of other factors, such as the variation of the CSI and limitation of the computational ability of the allocator, the effective duration of one particular DPM is assumed to be arbitrarily long so that the practical performance metrics converge to the expected values obtained by the queueing model and on the other hand, the requirement of computational ability to calculate the DPM can be ignored.

Table 3.1. Transmission modes with convolutionally coded modulation

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
Modulation	BPSK	QPSK	8-QAM	16-QAM	32-QAM	64-QAM
Coding rate R_c	1/2	1/2	3/4	9/16	3/4	3/4
Rate (bits/symbol)	0.5	1	1.5	2.25	3	4
a_n	274.7229	90.2514	67.6181	50.1222	53.3987	35.3508
g_n	7.9932	3.4998	1.6883	0.6644	0.3756	0.09
Γ_n(dB)	-1.5331	1.0942	3.9722	7.7021	10.2488	15.9784

3.3.3.3 Data transmission

In the data transmission part, AMC and truncated ARQ are applied by the SU system. The j-th PU channel is divided into N^j_{SNR} states according to the SNR at the SU receiver. Let $c^j(t) \in \{1,2,\cdots,N^j_{\text{SNR}}\}$ represent the channel condition state of the j-th channel at the t-th time slot. The evolution of the j-th channel condition states are modeled as a Markov chain using the same idea in [70], which is described by a N^j_{SNR}-by-N^j_{SNR} transition matrix $\boldsymbol{P}^j_{\text{CS}}$. According to the modulation method, the data transmission rate can be obtained from the channel state as explained below.

On the physical layer, convolutionally coded M_n-ary rectangular or square QAM modes is assumed to be used. These modulation schemes are adopted from the HIPERLAN/2 or IEEE 802.11a standards [54]. The details of the AMC scheme are shown in Table 3.1. The states of the j-th channel condition are divided into N^j_{SNR} parts. Different modulation schemes are used in each part to make the packet error rates of each part all equal to a required packet error rate P_{target}. Thus, the state boundaries in terms of SNR can be written as:

$$\Gamma_n = \frac{1}{g_n}\ln\left(\frac{a_n}{P_{\text{target}}}\right), \tag{3.5}$$

where $\Gamma_n \in \{\Gamma_0,\Gamma_1,\ldots,\Gamma_{N^j_{\text{SNR}}}\}$ denotes the boundary SNR of the n-th state. Then, by setting $\Gamma_0 = 0$ and $\Gamma_{N^j_{\text{SNR}}+1} = \infty$, the $N^j_{\text{SNR}}+1$ SNR boundaries of total N^j_{SNR} states can be obtained. When the SNR at an SU's receiver satisfies $\gamma \in (\Gamma_0,\Gamma_1)$, the SU does not transmit any packets due to the low SNR and when the SNR satisfies $\gamma \in (\Gamma_n,\Gamma_{n+1}), n \in \{1,2,\ldots,N^j_{\text{SNR}}-1\}$, mode n is used in the SU transmission.

Then, it can be determined how the channel condition evolves from time slot to time slot. It is assumed that the channel is Rayleigh: the SNR is exponentially distributed with the probability density function:

$$p(\gamma_i^j) = \frac{1}{\bar{\gamma}_i^j} \exp\left(-\frac{\gamma_i^j}{\bar{\gamma}_i^j}\right), \quad \gamma_i^j \geq 0, \tag{3.6}$$

where $\bar{\gamma}_i^j$ is the average SNR, and γ_i^j is the received SNR of the i-th SU through the j-th channel. The steady state probability of i-th SU on the j-th channel in state k can be calculated as given below:

$$\pi_{\text{CSi}}^j(k) = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma_i^j)\, \mathrm{d}x. \tag{3.7}$$

Let $\boldsymbol{P}_{\text{CSi}}^j = \begin{pmatrix} p_{0,0} & \cdots & p_{0,N_{\text{SNR}}^j-1} \\ \vdots & \ddots & \vdots \\ p_{N_{\text{SNR}}^j-1,0} & \cdots & p_{N_{\text{SNR}}^j-1,N_{\text{SNR}}^j-1} \end{pmatrix}$ denote the state transition probability matrix of the i-th SU on the j-th channel, where $p_{x,y}$ denotes the transition probability to the state y from the state x. Based on the same idea in [67], the transition probabilities can be obtained by:

$$p_{k,k+1} \approx \frac{f_m T_p}{\pi_{\text{CSi}}^j(k)} \sqrt{\frac{2\pi\Gamma_{k+1}}{\bar{\gamma}_i^j}} \exp\left(-\frac{\Gamma_{k+1}}{\bar{\gamma}_i^j}\right), \tag{3.8}$$

$$p_{k,k-1} \approx \frac{f_m T_p}{\pi_{\text{CSi}}^j(k)} \sqrt{\frac{2\pi\Gamma_k}{\bar{\gamma}_i^j}} \exp\left(-\frac{\Gamma_k}{\bar{\gamma}_i^j}\right), \tag{3.9}$$

where f_m is the maximum Doppler frequency. The transition is assumed to occur only between adjacent states: $p_{j,k} = 0$, for any $|j-k| > 1$, and $p_{k,k} = 1 - p_{k,k+1} - p_{k,k-1}$. Thus, using (3.5)-(3.9), the transition probability matrix of the entire system, $\boldsymbol{P}_{\text{CS}}^j$, is obtained. Let $V_{\text{rate}}(k)$ represent the number of packets that can be transmitted in one time slot when the channel condition state is k. According to the rate in Table 3.1, $\{V_{\text{rate}}(k), k \in \{0, 1, \cdots, 6\}\} = \{0, 2, 4, 6, 9, 12, 16\}$.

For the error recovery method, a truncated ARQ scheme is applied. This process is shown in Figure 3.3. If a packet is transmitted successfully, the transmitter will receive an ACK message from the receiver, otherwise, the receiver will return a negative-acknowledgment message, and the transmitter will transmit the packet again. If a packet has been transmitted unsuccessfully for L consecutive times, the transmitter will drop the packet, which is defined as a “drop” packet, then start to transmit the next packet. During the retransmission of the packets, specific technologies can be applied to improve the packet error rate, such as maximal ratio combining (MRC) [112]. In the model of the ARQ process, the packet error rate of the transmission and retransmissions can be configured according to the applied MRC method. To make the description of the proposed model concise, no further discussion on MRC

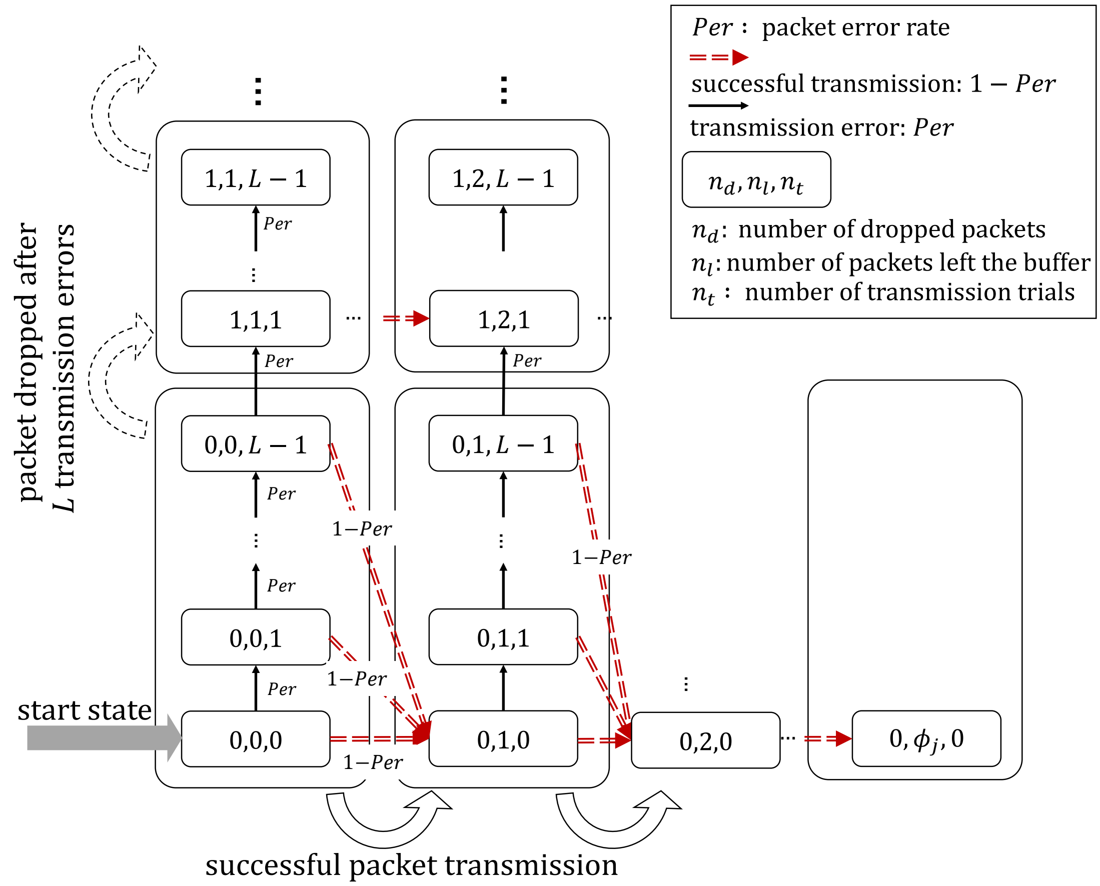


Figure 3.3. The truncated ARQ model

is carried out, and in the next part of this chapter, the packet error rates are assumed to be identical during the transmissions and retransmissions.

After data transmission, all packets arriving within the duration of the time slot will enter the buffer. If the total number of packets exceeds the buffer size, the overflow part of the packets will be rejected.

3.3.4 Formulation of the queueing model

In order to confine the size of the proposed model, the following analytical model is built on one selected SU (i-th). The procedure to analyze the other SUs in the system with the proposed model is identical.

3.3.4.1 Arrival process

It is assumed that the number of arrival packets at each SU during one time slot follows a Batch Bernoulli process, which can be found in the data collection from a wireless sensor network: for example, a client collects data from different sensors that broadcast their data. Data fusion of the obtained data from an arbitrary set of different sensors using a fusion algorithm can generate a data flow that follows a Batch Bernoulli process [113]. Let $\alpha_i(j)$ represent the probability that j packets arrive in one time slot at the i-th SU, and v_i is the maximum number of packets that arrive at the i-th SU in one time slot. Thus, the process can be described by a probability vector: $\boldsymbol{\alpha}_i = \{\alpha_i(0), \alpha_i(1), \cdots, \alpha_i(v_i)\}$.

3.3.4.2 Joint system states

Besides the arrival packets, the SU transmission is affected by the channel assignment result and the transmission scheme.

1 CHANNEL ASSIGNMENT STATE The channel assignment state is used to describe whether the SU can transmit on the PU channels. Let $d_i^j(t)$ represent the assignment result. $d_i^j(t) = 1$ represents that at the t-th time slot, the j-th channel is assigned to the i-th SU, while $d_i^j(t) = 0$ represents the opposite. Thus, a vector $\boldsymbol{d}_i^j(t) = \{\Pr(d_i^j(t) = 0), \Pr(d_i^j(t) = 1)\}$ can describe the entire possibility of channel assignment state. Let the matrix $\boldsymbol{P}_{\hat{O}}$ describe the relationship between the actual PU occupancy state and SU sensing result:

$$\boldsymbol{P}_{\hat{O}} = \begin{pmatrix} 1 - P_{\text{md}} & P_{\text{md}} \\ P_{\text{fa}} & 1 - P_{\text{fa}} \end{pmatrix}. \tag{3.10}$$

Let $\boldsymbol{Pd}_i^j$ describe the relationship between the SU sensing result and the channel assignment result:

$$\boldsymbol{Pd}_i^j = \begin{pmatrix} 1 & 0 \\ 1 - d_i^j & d_i^j \end{pmatrix}, \tag{3.11}$$

where d_i^j is the probability that the j-th channel is assigned to the i-th SU in Equation (3.4). Using Equation (3.10) and Equation (3.11), it can be determined that:

$$\boldsymbol{d}_i^j(t+1) = \boldsymbol{d}_i^j(t) \left(\boldsymbol{P}_{\hat{O}} \boldsymbol{Pd}_i^j\right)^{-1} \boldsymbol{P}_{\text{PU}}^j \left(\boldsymbol{P}_{\hat{O}} \boldsymbol{Pd}_i^j\right), \tag{3.12}$$

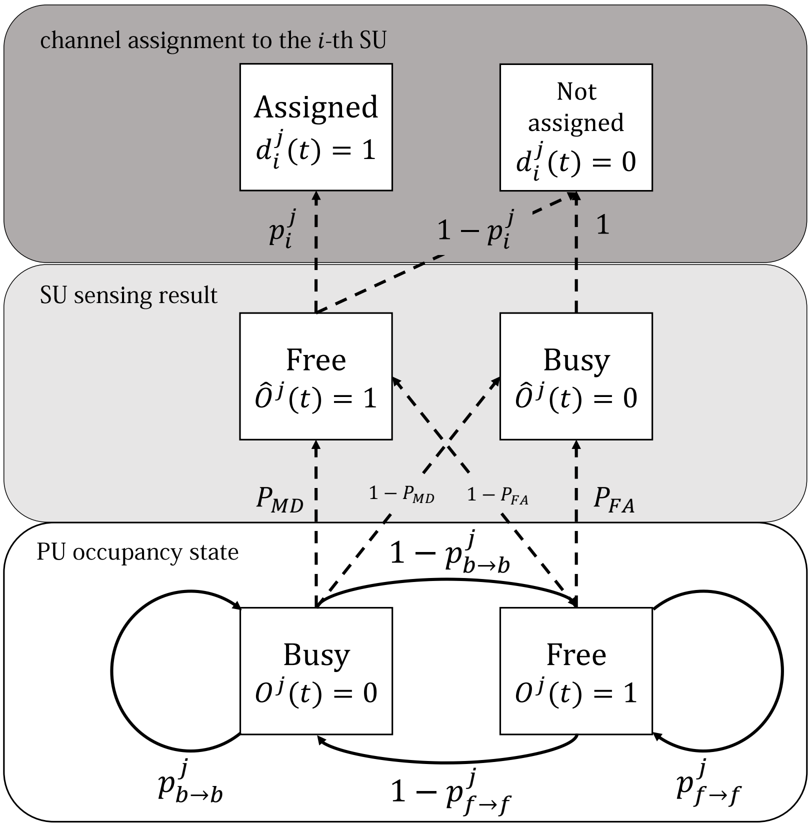


Figure 3.4. Channel assignment state.

from which, the assignment state of the SU as a Markov chain with the transition matrix can be obtained, as indicated below.

$$\boldsymbol{P}_d = \left(\boldsymbol{P}_{\hat{O}}\boldsymbol{Pd}_i^j\right)^{-1}\boldsymbol{P}_{\mathrm{PU}}^j\left(\boldsymbol{P}_{\hat{O}}\boldsymbol{Pd}_i^j\right). \tag{3.13}$$

The relationship among the PU occupancy state, SU sensing results and channel assignment results is shown in Figure 3.4.

2 SETUP OF THE JOINT STATE The state of one PU channel can be described by the assignment state and the channel condition state: $\{d_i^j(t), c_i^j(t) | d_i^j(t) \in \{0,1\}, c_i^j(t) \in \{1,2,\cdots,N_{\mathrm{SNR}}^j\}\}$. Let $\boldsymbol{S}_i(t)$ represent the joint state of channel condition of all M PU channels from the i-th SU's perspective:

$$\begin{aligned}\boldsymbol{S}_i(t) \triangleq \{&\left(d_i^1(t), c_i^1(t)\right), \left(d_i^2(t), c_i^2(t)\right), \ldots, \left(d_i^M(t), c_i^M(t)\right) | \\ &d_i^j(t) \in \{0,1\}, c_i^j(t) \in \{1,2,\cdots,N_{\mathrm{SNR}}^j\}, j \in \{1,2,\cdots,M\}\}.\end{aligned} \tag{3.14}$$

It is assumed that all PU channel conditions to SU are independent. According to the DPM and the activity of the PU, the allocation state of PU channels is also independent. Hence, the number of all channel condition states of M channels N_{all} is:

$$N_{\text{all}} = 2^M \cdot \prod_{j=1}^{M} N_{\text{SNR}}^j. \tag{3.15}$$

An integer index $s_i(t)$ is set up to label each state in $\boldsymbol{S}_i(t)$:

$$s_i(t) = \sum_{j=1}^{M} \left\{ \left[d_i^j(t) \cdot N_{\text{SNR}}^j + c_i^j(t) - 1 \right] \cdot \prod_{l=1}^{j} N_{\text{SNR}}^{l-1} \right\}, \tag{3.16}$$

where $N_{\text{SNR}}^0 = 0$ is set to make the expression concise. Then the transition probability matrix of the channel condition state can be obtained in which the states are sorted according to the indexes:

$$\boldsymbol{A} = \left(\boldsymbol{P}_d^1 \otimes \boldsymbol{P}_{\text{cs}}^1 \right) \otimes \cdots \otimes \left(\boldsymbol{P}_d^M \otimes \boldsymbol{P}_{\text{cs}}^M \right), \tag{3.17}$$

where $\otimes$ denotes the Kronecker product.

3.3.4.3 Service process

Now, the detailed transmission process under any given system state $s(t) \in \boldsymbol{S}_i(t)$ can be derived. As mentioned before, the truncated ARQ scheme for error correction and AMC on transmission is applied, thus the distribution of the number of transmitted packets from the channel condition state can be obtained. Let Φ_{all} represent the maximum number of packets that can be transmitted during one time slot. A series of functions $\boldsymbol{\Theta}$ is used to describe the relationship between packets that can be transmitted and the channel state:

$$\Theta(s(t)) = \left\{ \Theta_0\left(s(t)\right), \Theta_1\left(s(t)\right), \cdots, \Theta_{\Phi_{\text{all}}}\left(s(t)\right) \right\}, \tag{3.18}$$

where $\Theta_i\left(s(t)\right)$ represents the probability that i packets can be transmitted successfully when the channel state is $s(t)$. The derivation of $\boldsymbol{\Theta}$ is as given below.

1 THE DERIVATION OF $\boldsymbol{\Theta}$ According to the AMC settings, there are in total ϕ_j transmission attempts in one SU time slot from the j-th channel. Let $n_l \in \{0, 1, \ldots, \phi_j\}$ denote the number of packets leaving the queue from the channel, $n_t \in \{0, 1, \ldots, L-1\}$ denote the number of transmission trials and $n_d \in \left\{0, 1, \ldots, \left\lfloor \frac{\phi_j}{L} \right\rfloor \right\}$ denote the number of packets dropped caused by full SU buffer. Then, a state space $\{n_d, n_l, n_t\}$ can be built, which is called the ARQ state, and the transition matrix $\boldsymbol{P}_{\text{ARQ}}$ can be derived. The transition matrix $\boldsymbol{P}_{\text{ARQ}}$ consists of two levels of matrix block. The block of the first level describes the transition between number of transmission trials. $\boldsymbol{P}_0$ describes state transition when the transmission fails:

$$\boldsymbol{P}_0 = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ L-2 \\ L-1 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & \dots & L-1 \end{array} \\ \begin{pmatrix} 0 & \text{Per} & & & \\ & 0 & \text{Per} & & \\ & & \ddots & \ddots & \\ & & & 0 & \text{Per} \\ & & & 0 & 0 \end{pmatrix} \end{array}. \tag{3.19}$$

$\boldsymbol{P}_1$ describes the state transition when the transmission is successful:

$$\boldsymbol{P}_1 = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ L-1 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \dots & L-1 \end{array} \\ \begin{pmatrix} 1-\text{Per} & 0 & \dots & 0 \\ 1-\text{Per} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 1-\text{Per} & 0 & \dots & 0 \end{pmatrix} \end{array}. \tag{3.20}$$

$\boldsymbol{P}_{\text{drop}}$ describes the state transition when a packet is dropped after consecutive L failed transmission attempts:

$$\boldsymbol{P}_{\text{drop}} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ L-1 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \dots & L-1 \end{array} \\ \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \text{Per} & 0 & \dots & 0 \end{pmatrix} \end{array}. \tag{3.21}$$

Here, if the MRC method is applied, the Per can be modified accordingly to adapt the improvement on packet error rate during packet retransmissions.

The block of the second level describes the transition between number of departure packets. $\boldsymbol{Q}_i$ describes the transition between the number of departure packets when i packets have been dropped:

$$\boldsymbol{Q}_0 = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ \phi_j - 1 \\ \phi_j \end{array} \!\!\! \begin{array}{c} \begin{array}{ccccc} 0 & 1 & \dots & \phi_j - 1 & \phi_j \end{array} \\ \begin{pmatrix} \boldsymbol{P}_0 & \boldsymbol{P}_1 & & & \\ 0 & \boldsymbol{P}_0 & \boldsymbol{P}_1 & & \\ & & \ddots & \ddots & \\ & & & \boldsymbol{P}_0 & \boldsymbol{P}_1 \\ & & & 0 & \boldsymbol{P}_0 \end{pmatrix} \end{array}, \tag{3.22}$$

$$\boldsymbol{Q}_i = \begin{array}{c} \\ 0 \\ \vdots \\ i \\ \vdots \\ \phi_j - 1 \\ \phi_j \end{array} \!\!\! \begin{array}{c} \begin{array}{cccccc} 0 & \dots & i & \dots & \phi_j - 1 & \phi_j \end{array} \\ \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & \boldsymbol{P}_0 & \boldsymbol{P}_1 & & \\ & & & \ddots & \ddots & \\ & & & & \boldsymbol{P}_0 & \boldsymbol{P}_1 \\ & & & & 0 & \boldsymbol{P}_0 \end{pmatrix} \end{array}, \tag{3.23}$$

$\boldsymbol{D}_i$ describes the transition when a new packet is dropped when i packets have been dropped:

$$\boldsymbol{D}_0 = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ \phi_j - 1 \\ \phi_j \end{array} \!\!\! \begin{array}{c} \begin{array}{ccccc} 0 & 1 & \dots & \phi_j - 1 & \phi_j \end{array} \\ \begin{pmatrix} 0 & \boldsymbol{P}_{\text{drop}} & & & \\ & 0 & \boldsymbol{P}_{\text{drop}} & & \\ & & \ddots & \ddots & \\ & & & 0 & \boldsymbol{P}_{\text{drop}} \\ & & & 0 & 0 \end{pmatrix} \end{array}, \tag{3.24}$$

$$\boldsymbol{D}_i = \begin{array}{c} \\ 0 \\ \vdots \\ i \\ i+1 \\ \vdots \\ \phi_j-1 \\ \phi_j \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & \dots & i & i+1 & \dots & \phi_j-1 & \phi_j \end{array} \\ \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \boldsymbol{P}_{drop} & & & \\ & & & 0 & \boldsymbol{P}_{drop} & & \\ & & & & \ddots & \ddots & \\ & & & & & 0 & \boldsymbol{P}_{drop} \\ & . & & & & 0 & 0 \end{pmatrix} \end{array}. \tag{3.25}$$

Finally, the transition matrix can be obtained as below.

$$\boldsymbol{P}_{\text{ARQ}} = \begin{pmatrix} \boldsymbol{Q}_0 & \boldsymbol{D}_0 & & & \\ & \boldsymbol{Q}_1 & \boldsymbol{D}_1 & & \\ & & \ddots & \ddots & \\ & & & \boldsymbol{Q}_{\left\lfloor \frac{\phi_j}{L} \right\rfloor -1} & \boldsymbol{D}_{\left\lfloor \frac{\phi_j}{L} \right\rfloor -1} \\ & & & & \boldsymbol{Q}_{\left\lfloor \frac{\phi_j}{L} \right\rfloor} \end{pmatrix}. \tag{3.26}$$

Let $\Phi = \left\lfloor \frac{\phi_S}{L} \right\rfloor$, $\boldsymbol{P}_{\text{ARQ}}$ is a $((\Phi+1)\cdot(\phi_j+1)\cdot L) \times ((\Phi+1)\cdot(\phi_j+1)\cdot L)$ matrix. The transition diagram is shown in Figure 3.3. A series of 1-by-$((\Phi+1)\cdot(\phi_j+1)\cdot L)$ vector $\boldsymbol{b}_i$ is used to represent the probability distribution of the ARQ state $\{n_d, n_l, n_t\}$ after i transmissions. Let $b_i(I(n_d, n_l, n_t))$ denote the $I(n_d, n_l, n_t)$-th element of $\boldsymbol{b}_i$, which indicates the probability that the ARQ state is $\{n_d, n_l, n_t\}$:

$$I(n_d, n_l, n_t) = n_t + n_l \cdot L + n_d \cdot L(\phi_j + 1) + 1. \tag{3.27}$$

At the beginning of the transmission, the ARQ state starts at $\{0,0,0\}$, which is represented by $\boldsymbol{b}_0 = \{1,0,0,\dots\}$. Then $\boldsymbol{b}_i$, which represents the distribution of ARQ state after i transmissions, can be obtained as:

$$\boldsymbol{b}_i = \boldsymbol{b}_0 \cdot (\boldsymbol{P}_{\text{ARQ}})^i. \tag{3.28}$$

From $\boldsymbol{b}_i$ and Equation (3.27), the distribution of the number of packets leave the buffer and the number of packets dropped can be obtained:

$$\psi_m(i) = \sum_{j=(m-1)L(\phi_j+1)+1}^{mL(\phi_j+1)} b_i(j), 0 \leq m \leq \Phi, \tag{3.29}$$

$$\theta_m(i) = \sum_{a=0}^{\Phi} \left(\sum_{j=aL(\phi_j+1)+mL+1}^{aL(\phi_j+1)+(m+1)L+1} b_i(j) \right), 0 \leq m \leq \phi_j, \tag{3.30}$$

where $\psi_m(i)$ is the probability that m packets are dropped after i transmissions and $\theta_m(i)$ is the probability of m packets leaving the queue after i transmissions.

With Equation (3.30), the distribution of the total number of departure packets transmitted through all M channels can be obtained given the joint states: $s(t)$. A series of functions C^j is set up, and using V_{rate}, Equation (3.16), the number of transmission on the j-th channel when the channel state is $s(t)$, $C^j(s(t))$ can be derived. Then, the distribution of departure packets can be obtained when the channel state is $s(t)$, $\boldsymbol{\Theta}(s(t))$:

$$\boldsymbol{\Theta}\big(s(t)\big) = \boldsymbol{\theta}^1(C^1(s(t))) \circ \boldsymbol{\theta}^2(C^2(s(t))) \circ \cdots \circ \boldsymbol{\theta}^M(C^M(s(t))), \tag{3.31}$$

where $\circ$ denotes the convolution of vectors. $\boldsymbol{\Theta}(s(t))$ has $(N_{\text{all}}+1)$ elements. Let $\Theta_i(s(t))$ denote the i-th element of $\boldsymbol{\Theta}(s(t))$, which is the probability of i departure packets given the joint state $s(t)$. The $\boldsymbol{\Theta}(s(t))$, given all possible N_{all} different $s(t)$, can be obtained.

2 PROBABILITY MATRIX FOR DEPARTURE PACKETS Using $\boldsymbol{\Theta}(s(t))$ obtained above, a matrix $\boldsymbol{\mu}_i$, which represents the probability of i departure packets, given all possible channel states $s(t)$ in sequence, can be set up:

$$\boldsymbol{\mu}_i = \begin{pmatrix} \Theta_i\,(s(t)=1) \\ \vdots \\ \Theta_i\,(s(t)=N_{\text{all}}) \end{pmatrix} \otimes \mathbf{1}_{1\times N_{\text{all}}}, \tag{3.32}$$

where $\mathbf{1}_{1\times N_{\text{all}}}$ is a 1-by-N_{all} row vector whose elements are all 1.

For ease of describing the service process, a matrix $\boldsymbol{\mu}_i$ is set up, which is the same size as $\boldsymbol{A}$.

$$\boldsymbol{\mu}_i = \begin{pmatrix} \Theta_i\,(s(t)=1) \\ \vdots \\ \Theta_i\,(s(t)=N_{\text{all}}) \end{pmatrix} \otimes \mathbf{1}_{1\times N_{\text{all}}}, \tag{3.33}$$

where $\mathbf{1}_{1\times N_{\text{all}}}$ is a 1-by-N_{all} row vector whose elements are all 1.

3.3.4.4 Markov chain analysis

The following analysis is of one SU to keep the problem below a certain level of complexity without losing generality. If one only focuses on those packets that arrive at the transmitter and enter the buffer and are then transmitted, the system evolution from time slot to time slot can be treated as a discrete time Markov chain (DTMC). In the following part of the chapter, the analysis is of the i-th SU. The buffer size of the i-th SU is K_i. The DTMC state space of the system can be set up as indicated below.

$$\{(x_i(t), s_i(t)) | x_i(t) \in \{0, 1, \cdots, K_i\}; s(t) \in \{1, 2, \cdots, N_{\text{all}}\}\}, \tag{3.34}$$

where $x_i(t)$ is the number of packets in the i-th SU's buffer at the t-th time slot, $s_i(t)$ is the channel condition state of all M channels at the t-th time slot. Thus, the transition matrix of the system states, $\boldsymbol{P}$, can be obtained as shown in Equation (3.35).

$$
\boldsymbol{P}=
\begin{array}{c|c|ccccccc|cccccc|ccccc|c|ccc}
 & 0 & 1 & \ldots & v_i & v_i+1 & \ldots & N_{\text{all}}-1 & N_{\text{all}} & N_{\text{all}}+1 & \ldots & N_{\text{all}}+v_i & N_{\text{all}}+v_i+1 & \ldots & 2M\phi_S & 2N_{\text{all}} & \ldots & 2N_{\text{all}}+v_i+1 & \ldots & 3N_{\text{all}} & \ldots & \ldots & K_i-1 & K_i \\
0 & A_{(0,0)} & A_{(0,1)} & \cdots & A_{(0,v_i)} \\
\hline
1 & A_{(1,-1)} & A_{(1,0)} & & \cdots A_{(1,v_i-1)} & A_{v_i} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \ddots \\
v_i & A_{(v_i,-v_i)} & A_{(v_i,-v_i+1)} & \cdots & A_{(v_i,0)} & A_1 & \ddots \\
\vdots & \vdots & & & \vdots & \vdots & & \ddots \\
N_{\text{all}}-v_i+1 & \vdots & & & & & & \cdots A_{v_i-1} & A_{v_i} \\
N_{\text{all}}-v_i+2 & \vdots & & & & & & & A_{v_i-1} & A_{v_i} \\
\vdots & \vdots & & & & & & & & \ddots & \ddots \\
N_{\text{all}} & A_{(N_{\text{all}},-N_{\text{all}})} & \cdots & & & & & \cdots A_{-1} & A_0 & A_1 & \cdots & A_{v_i} \\
\hline
N_{\text{all}}+1 & & A_{-N_{\text{all}}} & \cdots & & & & \cdots & A_{-1} & A_0 & & \cdots A_{v_i-1} & A_{v_i} \\
\vdots & & & \ddots & & & & & \vdots & & & & & \ddots \\
2N_{\text{all}}-v_i+1 & & & & \ddots & & & & \vdots & & & & & & A_{v_i} \\
2N_{\text{all}}-v_i+2 & & & & & \ddots & & & \vdots & & & & & & A_{v_i-1} & A_{v_i} \\
\vdots & & & & & & \ddots & & \vdots & & & & & & \vdots & & \ddots \\
2N_{\text{all}} & & & & & & & & A_{-N_{\text{all}}} & A_{-N_{\text{all}}+1}\cdots & & & & & A_0 & A_1\cdots & & A_{v_i} \\
\hline
2N_{\text{all}}+1 & & & & & & & & & A_{-N_{\text{all}}}\cdots & & & & & A_1 & A_2\cdots & & & A_{v_i} \\
\vdots & & & & & & & & & & & \ddots & & & & & & \ddots & & & \ddots \\
\hline
\vdots & & & & & & & & & & & & & & & & & \ddots & & & \ddots \\
\hline
K_i-v_i & & & & & & & & & & & & & & & & & & & & & & \cdots A_{v_i-1} & A_{v_i} \\
K_i-v_i+1 & & & & & & & & & & & & & & & & & & & & & & \ddots A_{v_i-2} & A^{+}_{K_i-v_i+1} \\
\vdots & & & & & & & & & & & & & & & & & & & & & \ddots & \vdots & \vdots \\
K_i & & & & & & & & & & & & & & & & & & & & & & \ldots A_{-1} & A^{+}_{K_i}
\end{array}
\tag{3.35}
$$

.

The inner matrix blocks of $\boldsymbol{P}$ are defined as follows:

$$
\boldsymbol{A}_{(0,0)} = \alpha_i(0)\cdot\boldsymbol{A}, \quad \boldsymbol{A}_{(0,k)} = \alpha_i(k)\cdot\boldsymbol{A}, 1 \le k \le v_i, \tag{3.36}
$$

$$
\boldsymbol{A}_{\Delta k} = \left(\sum_{m=\max(\Delta k,0)}^{\min(\Delta k+N_{\text{all}},v_i)} \alpha_m \boldsymbol{\mu}_{(m-\Delta k)}\right) \circ \boldsymbol{A}, -N_{\text{all}} \le \Delta k \le v_i, \tag{3.37}
$$

$$
\begin{gathered}
\boldsymbol{A}^{+}_{k} = \left(\sum_{n=l-k}^{v_i} \alpha_n \left(\sum_{m=0}^{\max(k-l+n,0)} \boldsymbol{\mu}_m\right)\right) \circ \boldsymbol{A}, \\
K_i - v_i + 1 \le k \le K_i,
\end{gathered} \tag{3.38}
$$

where $\circ$ denotes the entrywise product. The other elements of $\boldsymbol{P}$ are all 0s.

3.3.5 Performance evaluation under DPM

3.3.5.1 Steady state distribution vector

The QR algorithm [114–116] can be applied to the transition matrix to find the steady state probability of $\boldsymbol{P}$, which is represented by a vector $\boldsymbol{\pi}$:

$$\boldsymbol{\pi} = \left\{\pi_1, \pi_2, \cdots, \pi_{(K_i+1)\times N_{\text{all}}}\right\}, \tag{3.39}$$

where π_i is the i-th element of $\boldsymbol{\pi}$. For easy evaluation of the performance measures, $\boldsymbol{\pi}$ is divided into K_i+1 vectors:

$$\boldsymbol{\pi} = \{\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_{K_i}\}, \tag{3.40}$$

where:

$$\boldsymbol{\pi}_i = \left\{\pi_{i\cdot N_{\text{all}}+1}, \pi_{i\cdot N_{\text{all}}+2}, \ldots, \pi_{(i+1)\cdot N_{\text{all}}}\right\}, 0 \le i \le K_i, \tag{3.41}$$

which indicates the probability of each joint channel state if there are i packets in the queue. Using the steady state distribution vector, it is able to determine some performance measures that are important. In order to evaluate the QoS of SUs, and furthermore to evaluate the performance of the DPM that is used to allocate the channels, the following performance measures are derived: average queue length, the packet rejection rate, gross throughput, the packet collision rate, the packet drop rate and average packet delay.

3.3.5.2 Queue length

Let q_l be the number of packets in queue. The probability of l packets in queue can be obtained by summing all elements in $\boldsymbol{\pi}_l$:

$$\Pr(q_l = l) = \boldsymbol{\pi}_l \mathbf{1} = \sum_{m=l\cdot N_{\text{all}}+1}^{(l+1)\cdot N_{\text{all}}} \pi_m. \tag{3.42}$$

Therefore, the average queue length Q_l can be obtained:

$$Q_l = \sum_{m=0}^{K} m \Pr(q_l = m). \tag{3.43}$$

3.3.5.3 Packet rejection rate

At the beginning of an SU time slot, if the number of arrival packets plus the number of packets in queue after transmission exceeds the SU buffer size, the overflow packets are rejected. Let r_n denote the probability that n packets are rejected at the beginning of a time slot. r_n can be obtained as :

$$r_n = \sum_{m=n}^{v} \alpha_i(m) \left(\sum_{b=k-m+n}^{k} \boldsymbol{\pi}_b (\boldsymbol{\mu}_{b-k-n+m})^T \right). \tag{3.44}$$

Hence the expected number of rejected packets during one time slot is :

$$\mathrm{Rej} = \sum_{\mathrm{n=1}}^{\mathrm{v}} \mathrm{n} \cdot \mathrm{r_n}. \tag{3.45}$$

3.3.5.4 Gross throughput

Gross throughput is indicated by the number of packets that leave the queue during one time slot, no matter whether they are successfully received, in collision with PU or dropped. Let η denote the average gross throughput, which can be obtained as:

$$\begin{aligned} \eta &= \sum_{m=1}^{N_{\text{all}}-1} \left(\sum_{a=1}^{m-1} a\boldsymbol{\pi}_m \boldsymbol{\mu}_a + m\boldsymbol{\pi}_m \sum_{b=m}^{N_{\text{all}}} \boldsymbol{\mu}_b \right) + \sum_{m=N_{\text{all}}}^{K} \sum_{a=1}^{N_{\text{all}}} a\boldsymbol{\pi}_m \boldsymbol{\mu}_a \\ &= \sum_{m=1}^{K} \sum_{a=1}^{N_{\text{all}}} \min(m,a) \boldsymbol{\pi}_m \boldsymbol{\mu}_a \end{aligned}. \tag{3.46}$$

3.3.5.5 Packet collision rate

When miss detection happens in one time slot, all the transmission packets will be in collision with the PU transmission. The expected number of collision packets can be derived as:

$$\mathrm{Col} = \frac{\mathrm{Pr(O^j(t)=0)P_{md}}}{\mathrm{Pr(O^j(t)=0)P_{md} + (1-Pr(O^j(t)=0))(1-P_{fa})}} \cdot \eta, \tag{3.47}$$

where $\Pr(O^j(t)=0) = \frac{p^i_{\text{b}\to\text{b}}}{p^i_{\text{b}\to\text{b}} + p^i_{\text{f}\to\text{f}}}$ is the probability of a channel being occupied by PU.

3.3.5.6 Packet drop rate

In order to derive the expected number of dropped packets, it is necessary to calculate the percentage of dropped packets of all departure packets. From Equation (3.7), the steady distribution of channel condition state of the j-th channel $\boldsymbol{\pi}_{\mathrm{csi}}^{j}$ can be obtained. When the channel is in state k, the expectation of expected number of dropped packets through channel j is:

$$E_{n_d}^{j}(i)=\sum_{a=1}^{N_{\mathrm{all}}}\left(a\cdot\sum_{m=a\phi_jL}^{(a+1)\phi_jL-1}\psi_m(V_{\mathrm{rate}}(k))\right), \tag{3.48}$$

where $\psi_m(V_{\mathrm{rate}}(k))$ represents the probability of m dropped packets under ARQ state k. The expected number of packets departing through channel j is:

$$E_{n_l}^{j}(k)=\sum_{m=1}^{\phi_j}m\cdot\theta_m(V_{\mathrm{rate}}(k)). \tag{3.49}$$

The details of $\psi_m(V_{\mathrm{rate}}(k))$ and $\theta_m(V_{\mathrm{rate}}(k))$ can be found in Equation (3.29) and Equation (3.30). Hence the packet drop rate is:

$$\rho_{\mathrm{drop}}=\sum_{j=1}^{M}p_i^j\frac{\sum_{m=1}^{N_{\mathrm{SNR}}^j}E_{n_d}^{j}(m)\pi_{\mathrm{csi}}^{j}(m)}{\sum_{m=1}^{N_{\mathrm{SNR}}^j}E_{n_l}^{j}(m)\pi_{\mathrm{csi}}^{j}(m)}, \tag{3.50}$$

where p_i^j is the probability of the j-th channel allocated to the i-th SU. The expectation of number of dropped packets during one time slot is:

$$\mathrm{Drop}=\rho_{\mathrm{drop}}\cdot\eta. \tag{3.51}$$

3.3.5.7 Average packet delay

The average packet delay in this chapter is defined as the expected time interval between the arrival of a packet and the start of its transmission. The average packet delay of a tagged packet consists of two parts: the transmission time of the packets that are in the buffer before the packet arrival process and the transmission time of the packets that arrive in the same time slot but are transmitted before the tagged packet. Thus, the average packet delay can be calculated as:

$$T_w=\frac{Q_l+\frac{\sum_{m=1}^{v_i}m\cdot\alpha_i(m)}{2}}{\eta}. \tag{3.52}$$

3.3.6 Numerical results and discussion

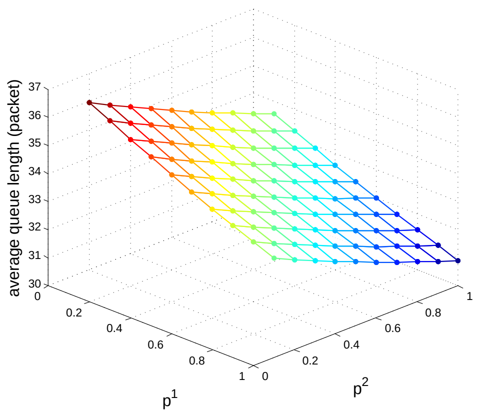
Using the proposed queueing model, several selected performance measures of one SU in the multi-channel multi-user CRN are evaluated. If not mentioned particularly, values of the parameters are assumed as follows:

- Number of SU: 2; Number of PU channels: 2
- Packet length: 2 ms, carrier frequency: 5 GHz, channel condition settings are in Section 3.3.3.3. Target error rate: 1%
- PU activity settings: $p^1_{\mathrm{f}\to\mathrm{f}} = p^2_{\mathrm{f}\to\mathrm{f}} = p^1_{\mathrm{b}\to\mathrm{b}} = p^2_{\mathrm{b}\to\mathrm{b}} = 0.5$
- Spectrum sensing settings: $P_{\mathrm{fa}} = 0.00661$, $P_{\mathrm{md}} = 0.00661$
- SU buffer size: $K_1 = 60$. $\boldsymbol{\alpha}_i = \{0.2, 0.2, 0.2, 0.2, 0.2\}$.
- Average SNR of the 2 channels are 10 dB.

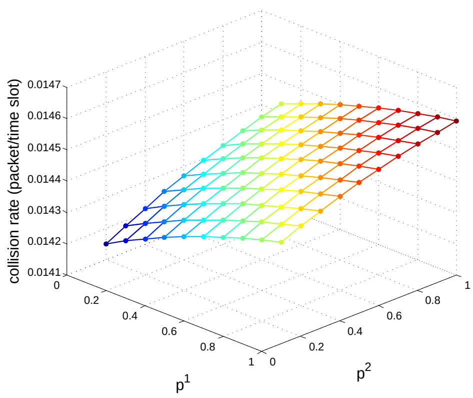
The impact of DPM on performance measures and the impact of different environmental parameters and SU system settings on performance measures are investigated from the numerical results.

3.3.6.1 Impact of DPM on performance measures

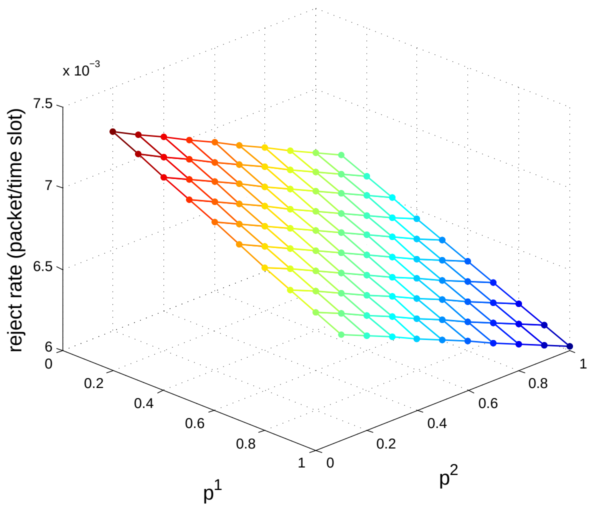
In the proposed model, the capacities of the PU channels are much more adequate for the traffic of SUs, so the author focuses on the measures that describe the data transmission failures and the demands of system settings. Figure 3.5 shows how average number of packets in a queue (Figure 3.5(a)), the number of collision packets (Figure 3.5(b)) and number of rejected packets (Figure 3.5(c)) changes under different distribution probabilities of the SU. It is clear and intuitive that if the SU gets more opportunity to be allocated to PU channels, the average number of packets in the buffer and the number of rejected packets will decrease, and meanwhile the number of collision packets will increase because of the high throughput. The relationship shown in the results obtained by the proposed framework may be used in the following ways:



(a) Average queue length versus DPM.



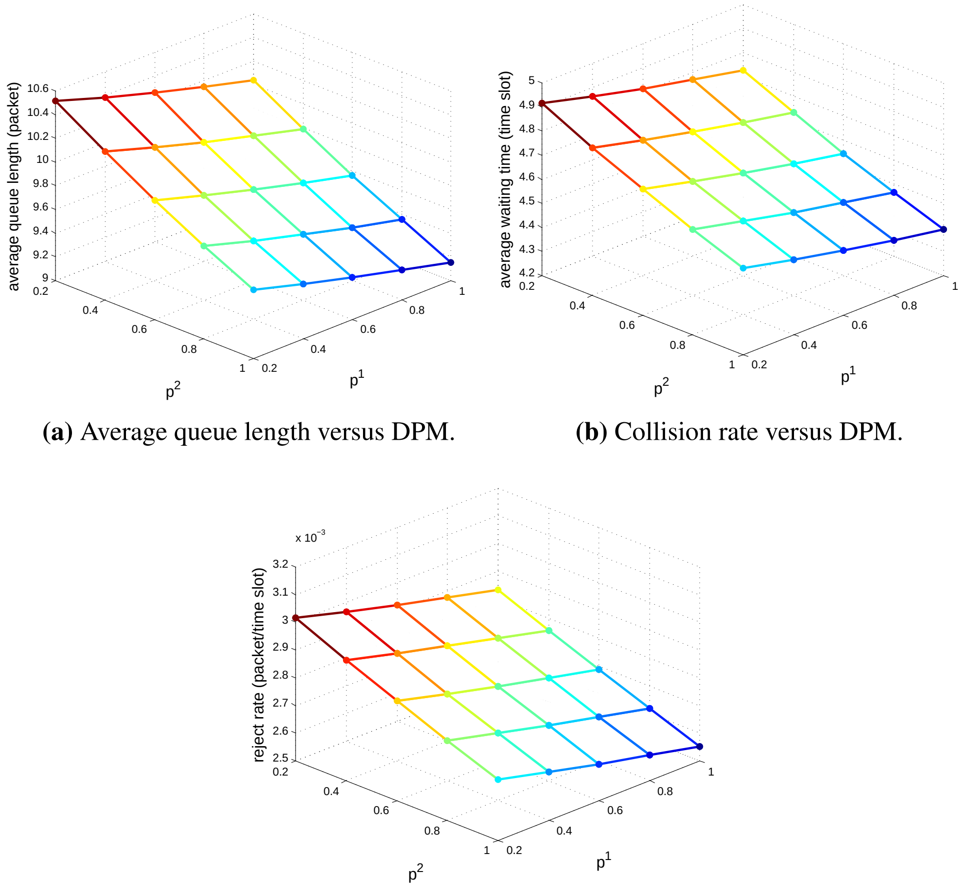
(b) Collision rate versus DPM.



(c) Rejection rate versus DPM.

Figure 3.5. Performance measures on different DPM. p^1 and p^2 are the distribution probabilities of the SU to channel 1 and channel 2, respectively.

- When designing a CR system, if there are some limits on certain performance measures (e.g. the data rate must be larger than a certain standard), the boundary of distribution probability of the SU that the applied channel allocation protocol must guarantee can be obtained.
- With the help of the relationship between DPM and performance measures, optimization of one performance measure or some of the performance measures combined can be carried out easily.



(a) Average queue length versus DPM.

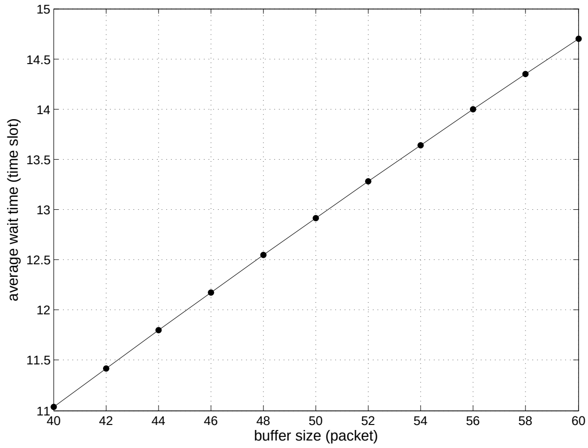
(b) Collision rate versus DPM.

(c) Rejection rate versus DPM.

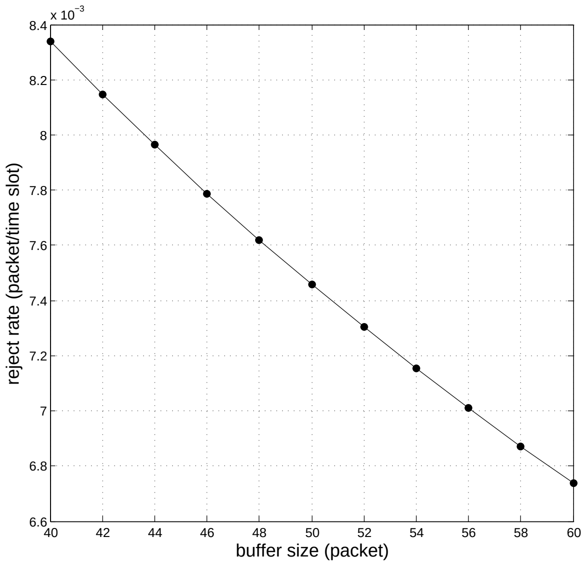
Figure 3.6. Performance measures on different SNR. The SNR of PU channel 1 is 10 dB, and PU channel 2 is 30 dB.

3.3.6.2 Impact of different environmental parameters on performance measures

In Figure 3.6, the channel settings are changed to show how the performance measures change under various DPMs. The average SNR at the SU receiver of PU channel 2 is set to 30 dB. It is evident that the changes in the distribution probability of that channel (p^2) have a greater influence on the performance measures, including average queue length, average waiting time and rejection rate. So, if the CR system is sharing different channels or even channels from different kinds of systems with different parameters, the DPM protocol can be adjusted accordingly to fulfill the target performance measures or optimize the performance as aforementioned in Section 3.3.6.1.



(a) Average wait time versus SU buffer size.



(b) Rejection rate versus SU buffer size.

Figure 3.7. Performance measures versus SU buffer size.

3.3.6.3 Impact of system settings on performance measures

In this part, the impact of the buffer size K_i and the truncated ARQ setting L on the performance measures is investigated to demonstrate how the proposed model can be used to help design CR system settings to meet the performance requirements. In Figure 3.7, the buffer size of the SU is set from 40 to 60 and changes in the performance measures are investigated under the DPM parameters: $p^1 = p^2 = 0.5$. In Figure 3.7(a), the average packet waiting time increases when the SU buffer size becomes a larger because larger buffer size will cause more packets to wait in the buffer. Figure 3.7(b)

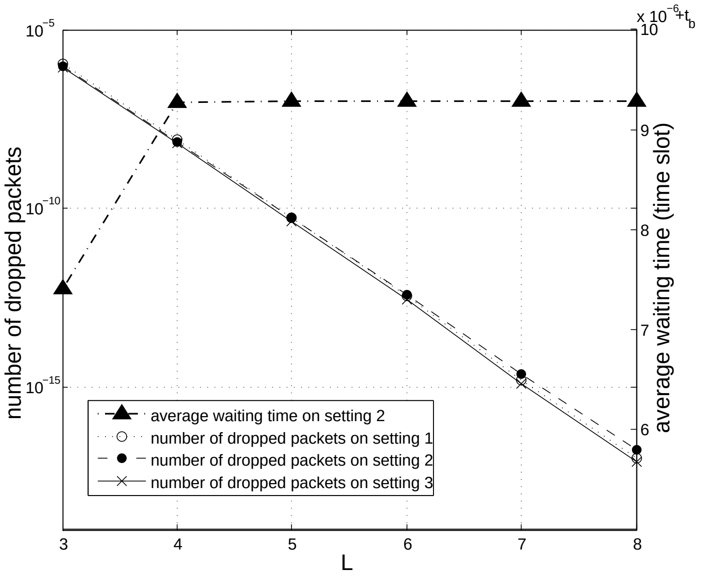


Figure 3.8. The impact of ARQ parameter L on performance measures under different system settings. $t_b = 11.03447$. Setting 1: SU packet arrival vector: $\boldsymbol{\alpha} = \{0.2, 0.2, 0.2, 0.2, 0.2\}$, SNR of PU channel 1 and 2: 10 dB. Setting 2: $\boldsymbol{\alpha} = \{0.2, 0.4, 0.4\}$, SNR of PU channel 1 and 2: 10 dB. Setting 3: $\boldsymbol{\alpha} = \{0.2, 0.4, 0.4\}$, SNR of PU channel 1 and 2: 30 dB.

shows that the number of rejected packets will decrease because the probability that the buffer becomes full is less when the buffer size is larger.

In Figure 3.8, the truncated ARQ parameter L is set from 3 to 8 and the impact on the number of dropped packets and average waiting time under different SU arrival patterns and channel conditions is investigated. The packet drop rate is affected mainly by the ARQ parameter L and almost unaffected by the arrival rate and the channel conditions. The packet average waiting time is increased by L when L is less than 5; after that there is no distinct change in the waiting time with the increase in L. To avoid the performance degradation such as the packet average waiting time in this research, L is recommended to be set around 3 in practice according to the numerical results in this section.

The relationship between the CR system settings and selected performance measures revealed by the proposed model can help analyze which CR system settings is the key factor influencing particular performance measures. Further, the CR system can be designed properly to meet the selected performance requirements.

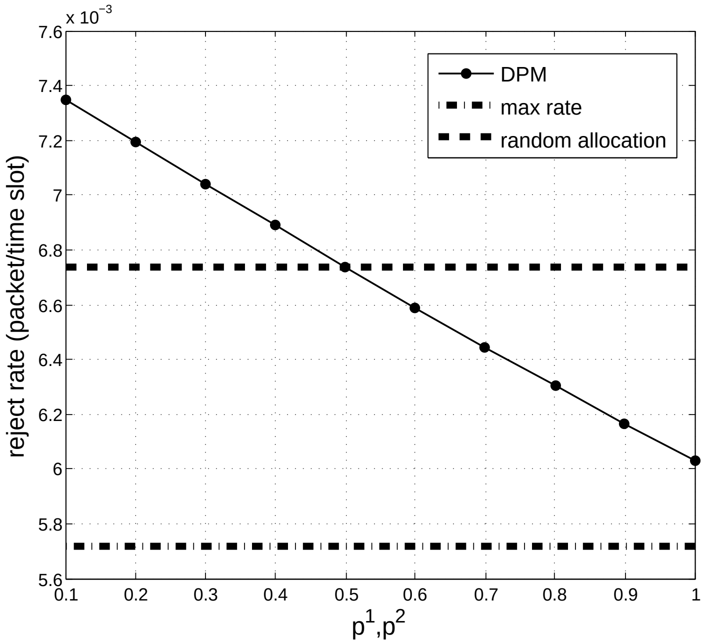


Figure 3.9. Comparison of average queue length and rejection rate of different channel allocation methods

3.3.7 Comparison of DPM and other protocols

In this section, the DPM, the maximum rate (MR) scheme [117, 118] and random allocation scheme are compared. The MR scheme in this chapter works as follows: When a PU channel is not occupied and is available to the SU system, the SU system will assign it to the SU that can transmit at the highest rate. In the random allocation scheme, all channels are allocated randomly to each SU with the same probability. In a CR system of N SUs and M channels, the random allocation scheme can be represented by setting the DPM parameter matrix as $d_i^j = 1/M$ for $i \in \{1,\ldots,N\}1, j \in \{1,\ldots,M\}$. The main difference in the methods is the way in which PU channels are allocated to the SUs after the SU system has collected all necessary information from the PU and SU systems. In order to compare the complexities of the allocation procedure under these methods, the time costs of channel allocation computation of the three channel allocation methods by MATLAB simulation are presented in Table 3.2. These time costs have been normalized to eliminate the influence caused by the selection among different computing environment (software and hardware) and represent the computing resource consumption during the channel allocation procedure in the CR system. When using DPM, the most computing resource consuming part is the configuration of the distribution matrix under various system settings and environmental parameters. However, this configuration can be done in advance rather

Table 3.2. Time Costs of Channel Allocation Computation

Parameters	Time cost ratio to DPM		
	DPM	MR	random allocation
number of SU: 2 number of channels: 2	1	10.8	0.24
number of SU: 2 number of channels: 10	1	11	0.22
number of SU: 4 number of channels: 10	1	9.1	0.26

than when the channel is actually allocated. Thus, no extra computing resource is requested during the channel allocation procedure. Then, the performance under different channel allocation methods is shown. Figure 3.9 shows the average number of rejected packets in one time slot when the DPM parameters are $p^1 = p^2$ and $p^1, p^2 \in \{0.1, 0.2, \ldots, 1\}$ and the comparison with the average number of rejected packets under the MR scheme and random allocation scheme. When the parameters are set properly ($p^1, p^2 > 0.5$), the DPM can outperform the random allocation. The optimal average number of rejected packets under the DPM is close to the number under the MR scheme. However, if the MR scheme is applied, it will cost the SU system more resources to compare the channel condition and the data rate of the SU mobile stations than DPM and random allocation as shown in Table 3.2. Furthermore, the accuracy of the results of the comparison will affect the practical performance. So when designing the channel allocation schemes, there is a trade-off between the complexity and the performance to be considered, and the proposed framework can provide references for that.

3.4 CONCLUSION

In this chapter, a channel allocation protocol creation mechanism, namely DPM, is introduced. The applications of the DPM are the creation of channel allocation protocols and analysis of channel allocation protocols. As the first part of the thesis, the DPM is applied as a channel allocation protocol in this chapter. The DPM as a configurable channel allocation protocol is able to describe the allocation result more comprehensively than other channel allocation protocols, such as opportunistic scheduling and random allocation. A queueing model to study how the channel allocation results under different environmental parameters and different CR system settings affect the system performance measures is established. Important performance measures are derived by fitting link adaptation technologies, such as AMC and truncated ARQ and DPM, into the queueing model. The influence of the CR system settings on performance measures in different channel allocation situations described by the DPM is shown. The comparison of channel allocation protocols shows that first, the DPM outperforms the random allocation protocol; second, the DPM is able to yield performance close to the MR protocol but with less cost of computational resources during the channel allocation process. This chapter showed the advantages of the DPM as a channel allocation protocol and the potential of the DPM to provide references to determine the SU system settings to achieve the performance requirements. The further application of DPM will be introduced in subsequent chapters.

CHAPTER 4 MAXIMUM THROUGHPUT CHANNEL ALLOCATION PROTOCOL BASED ON DPM

4.1 CHAPTER OVERVIEW

To fulfill the promise of improving radio resource efficiency, the channel allocation protocol plays an essential role in CRNs. To design an appropriate protocol, an estimation of the probable performance should be carried out in advance taking into consideration the RA objectives. However, in a multi-use multi-channel CRN, some of the performance metrics under a given channel allocation protocol are complex to analyze, especially when the channel allocation objectives are to achieve some overall goals, such as maximizing the throughput of all the users. In this chapter, a queueing analytical framework based on DPM to obtain the performance metrics of various types is introduced. Then an MT channel allocation protocol is built using the DPM mechanism and its performance is analyzed and compared with related protocols by the proposed framework. The novelty of the DPM mechanism to analyze multi-user multi-channel CRN and model various channel allocations is described in this chapter.

4.2 BACKGROUND

The application of queueing analysis in telecommunication research is well developed and proves to be effective and efficient for research on CR technology. Applying queueing analysis, various performance measures can be derived efficiently. Thus, the way in which different channel allocation protocols affect performance measures can be investigated, as well as how a channel allocation protocol should be designed to meet the requirements of performance in the CRN.

As introduced in Section 2.4, there is a group of protocols that make use of CSI aimed at achieving certain preset objectives. Among the objectives, achieving optimal throughput for the SU system is popular, for example through opportunistic spectrum scheduling [48] [119] and the MR scheme [120]. In [121], the authors take a further step to take the primary user's activity into account. However, the existing protocols have some limitations. Firstly, the criteria of these channel allocation protocols are only based on the transmission capacity through PU channels of each SU. However, the transmission demand of the SU is ignored. Secondly, the objective of most analysis is a single selected SU in the SU system and the behavior of other SUs is assumed to be independent and identical to that of the selected SU. The correlation of the behavior of the SUs is ignored and the performance evaluation of the allocation protocol is presented only by the performance of the selected SU rather than the performance of all SUs.

In this chapter, a new channel allocation protocol named the MT protocol to allocate multiple PU channels to multiple SUs is introduced. This allocation protocol is aimed at maximizing the overall throughput of all the SUs in the SU system, taking into consideration both the SU transmission capacities and the SU transmission demands. To apply the MT protocol to investigate its performance, a novel queueing analysis framework for multi-user multi-channel CRN is proposed. In this framework, the queueing model is adapted to the analysis on each individual SU in the system. All the SUs are modeled independently, so the overall performance of the SU system can be obtained accurately.

In the remainder of this chapter, the queueing analytical framework is introduced first. Then, the MT protocol is applied as well as an MR scheme from the related literature that works as a reference in this framework. Finally, comparisons of the protocols are discussed through numerical results and the novelty of the MT protocol and the framework is expounded.

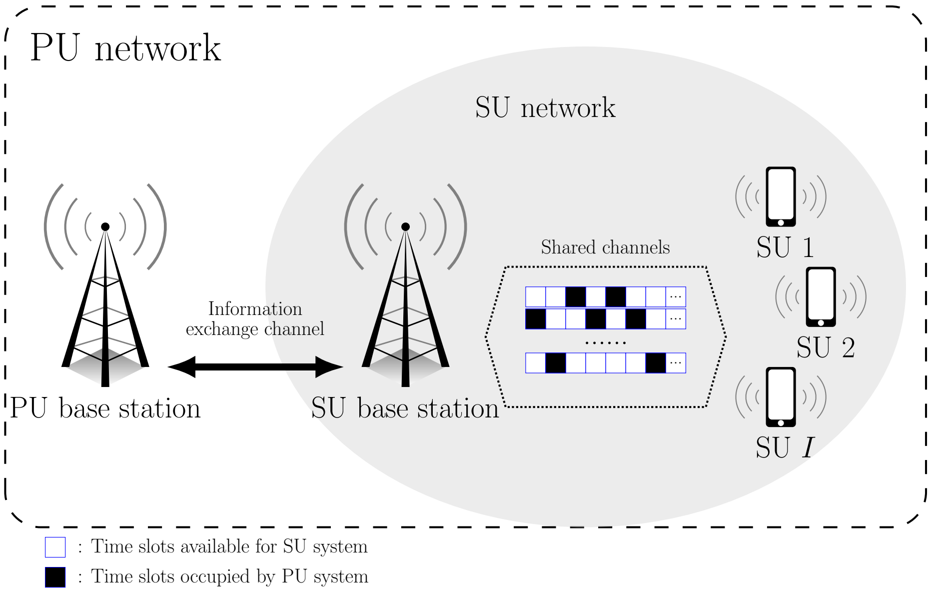


Figure 4.1. System model

4.3 FORMULATION OF THE QUEUEING MODEL FOR MULTIUSER MULTICHANNEL CRN

4.3.1 Basic model settings

In this chapter, the uplink channel allocation for an overlay [32] CRN with multiple SUs and multiple channels is considered. An example of the system is shown in Figure 4.1. A reliable CSI channel between the PU and SU system that can help exchange CSI is assumed. The CRN works within the coverage of one or more PU networks and comprises one SU base station and multiple SUs working with it. The number of SUs is M. The number of PU channels that are available to the SU system is J. The data transmissions in the system are in the form of discrete packets and the SU system works under the identical time slot structure synchronized with the PU system. The size of the buffer that stores the packets of each SU waiting to be transmitted is finite. The buffer size of the i-th SU is K_i.

The channel allocation protocol is the rule that the SU system uses to decide at the beginning of a time slot which SU can transmit on the channel where the PU is absent. The objective of the channel allocation protocol in CRN is to fulfill QoS requirements such as maximizing the throughput, minimizing the packet lost rate, and minimizing the packet delay.

4.3.2 Primary user activity model

The SU system cannot transmit on a channel while a PU is transmitting on it. The PU activities on a channel are modeled as a time-homogeneous first-order Markov process with two states. The process and the behavior of the SU system are independent. Let $O^j(t) \in \{0,1\}$ represent the PU activity state of the j-th channel at the t-th time slot. When a PU is transmitting, the state is called "busy" and $O^j(t) = 0$, otherwise the state is called "free" and $O^j(t) = 1$. The transition matrix of this Markov process is described below:

$$\boldsymbol{P}^j_{\mathrm{PU}} = \begin{pmatrix} p^j_{\mathrm{b}\to\mathrm{b}} & 1 - p^j_{\mathrm{b}\to\mathrm{b}} \\ 1 - p^j_{\mathrm{f}\to\mathrm{f}} & p^j_{\mathrm{f}\to\mathrm{f}} \end{pmatrix} \tag{4.1}$$

where $p^j_{\mathrm{b}\to\mathrm{b}} = \Pr(O^j(t) = 0 | O^j(t-1) = 0)$ denotes the probability that a PU is transmitting on the j-th channel (the channel state is "busy"), given that there was a PU transmitting on the channel (the channel state is "busy") at the previous time slot, and $p^j_{\mathrm{f}\to\mathrm{f}} = \Pr(O^j(t) = 1 | O^j(t-1) = 1)$ denotes the probability that the j-th channel state becomes "free" from the "free" state. Through the model above, the percentage of PU occupancy (the "busy" state) θ and average PU occupancy duration t_{PU} can be calculated as follows:

$$\theta = \frac{1 - p^j_{\mathrm{f}\to\mathrm{f}}}{2 - p^j_{\mathrm{b}\to\mathrm{b}} - p^j_{\mathrm{f}\to\mathrm{f}}} \tag{4.2}$$

$$t_{\mathrm{PU}} = \frac{1}{1 - p^j_{\mathrm{b}\to\mathrm{b}}} \tag{4.3}$$

In this model, the PU occupancy state of all J channels at the t-th time slot can be represented by a vector $\boldsymbol{o}$:

$$\boldsymbol{o} \in \mathcal{O} = \left\{ \left(O^1(t), O^2(t), \ldots, O^J(t)\right) | O^j(t) \in \{0,1\} \right\}. \tag{4.4}$$

4.3.3 Secondary system transmission model

The data transmission of the SU system is made under the structure of time slot synchronized with the PU system. A time slot is divided into three consecutive parts, namely the spectrum sensing part, the channel allocation part and the data transmission part.

4.3.3.1 Spectrum sensing

The SU system senses whether any PU is transmitting on every PU channel. At the end of the spectrum sensing part, the SU system makes an estimation of whether there is a PU transmitting on each channel and collects the information on the transmission conditions (e.g: the SNR at transceivers) of each SU mobile station. Here, in this chapter, the perfect spectrum sensing is assumed, which means the PU behavior (transmitting or not) on every channel will be detected correctly.

4.3.3.2 Channel allocation

Based on the information the SU system collects during the spectrum sensing, the SU system allocates the channels that are available to the SUs to transmit their data in the current time slot. Here, the following assumptions are made: One PU channel can be allocated to only one SU. One SU can transmit through multiple PU channels in one time slot. The DPM mechanism is applied here. It works as follows: A J-by-$(M+1)$ DPM is set up as indicated below.

$$\boldsymbol{P}_{\mathrm{DPM}} = \begin{pmatrix} d_0^1 \, d_1^1 \cdots d_M^1 \\ \vdots \;\; \vdots \;\; \ddots \;\; \vdots \\ d_0^J \, d_1^J \cdots d_M^J \end{pmatrix}, \tag{4.5}$$

where d_0^j represents the probability that the j-th channel is not allocated to any SU, and d_i^j for $i \neq 0$ represents the probability that the SU system allocates the j-th channel to the i-th SU. According to the assumptions, in $\boldsymbol{P}_{\mathrm{DPM}}$: $\sum_{i=0}^{M} d_i^j = 1$ for any $j \in [1, J]$. In order to investigate the optimal performance that a channel allocation protocol can achieve, the following assumptions are made: the computational ability of the channel allocator is qualified for the calculation of the DPM at the beginning of every time slot considering all the CSI and the information of every SU. Thus, the effective duration of DPM is one time slot.

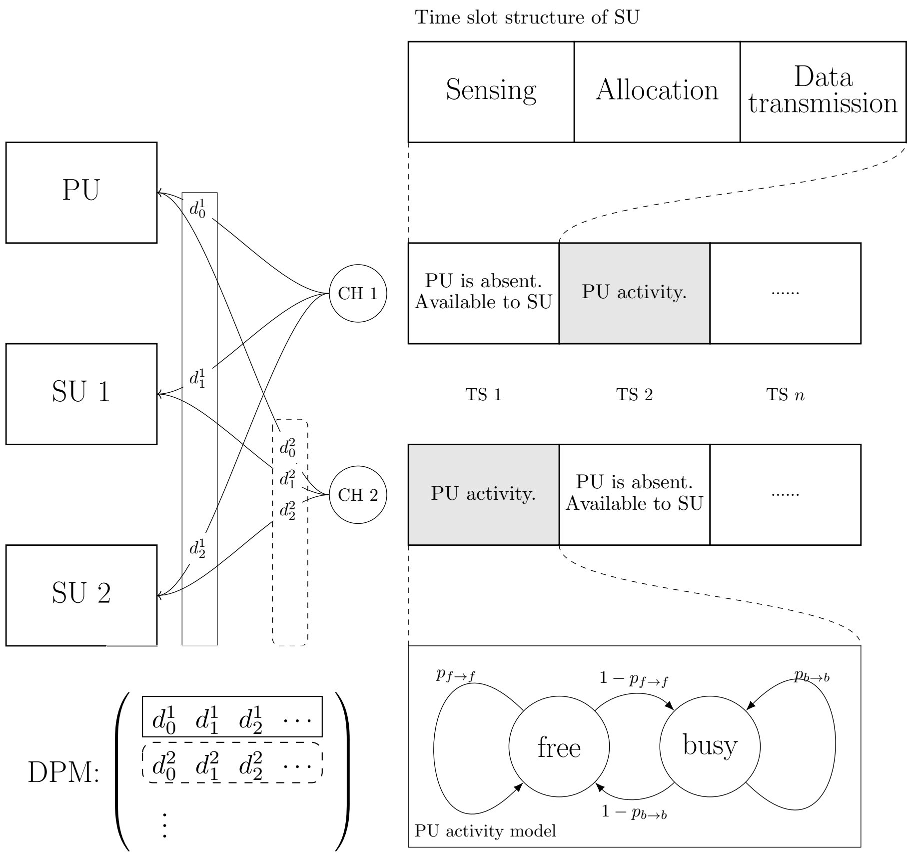


Figure 4.2. Illustration of DPM, time slot structure of SU system and the PU activity model.

4.3.3.3 Data transmission

The channel condition for SU data transmission is modeled as a Markov process. CSI exchange between the PU and SU system is assumed to be perfect. The AMC is applied in the SU system. The SNR of the j-th PU channel at the SU receiver is divided into N^j_{SNR} non-overlapped states. Let $c^j(t) \in \{1, 2, \cdots, N^j_{\mathrm{SNR}}\}$ represent the channel condition state of the j-th channel at the t-th time slot. The evolution of the j-th channel condition states is modeled as a Markov chain using a similar idea as in [70]. The model can be described by a N^j_{SNR}-by-N^j_{SNR} transition matrix $\boldsymbol{P}^j_{\mathrm{CS}}$, which is obtained by the procedure described below. The convolutionally coded M_n-ary rectangular and square QAM modes here are used as the modulation method adopted from the HIPERLAN/2 and IEEE 802.11a standards. The details of the AMC scheme are listed in Table 4.1.

Based on the SNR at the SU's transceiver, the j-th channel condition is divided into N^j_{SNR} states.

Table 4.1. Transmission modes with convolutionally coded modulation

	Mode 1	Mode 2	Mode 3
Modulation	BPSK	QPSK	8-QAM
Coding rate R_c	1/2	1/2	3/4
Rate(bits/symbol)	0.5	1	1.5
a_n	274.7229	90.2514	67.6181
g_n	7.9932	3.4998	1.6883
Γ_n(dB)	-1.5331	1.0942	3.9722

Different modulation schemes are applied to make the average packet error rates under each state all equal to a preset packet error rate requirement, P_{target}. Thus, the boundary of these states in terms of SNR can be obtained as described below [89].

$$\Gamma_n = \frac{1}{g_n} \ln\left(\frac{a_n}{P_{\text{target}}}\right), \tag{4.6}$$

where $\Gamma_n \in \{\Gamma_0, \Gamma_1, \dots, \Gamma_{N^j_{\text{SNR}}}\}$ denotes the boundary SNR of the n-th state. Then, by setting $\Gamma_0 = 0$ and $\Gamma_{N^j_{\text{SNR}}+1} = \infty$, the $N^j_{\text{SNR}}+1$ SNR boundaries of the N^j_{SNR} states can be obtained. When the SNR at an SU's transceiver satisfies the following condition: $\gamma \in (\Gamma_0, \Gamma_1)$, the SU does not try to transmit any packets due to the low SNR. When $\gamma \in (\Gamma_n, \Gamma_{n+1}), n \in \{1, 2, \dots, N^j_{\text{SNR}} - 1\}$ is satisfied, the SU transmits under the mode n.

Thus, the process of the channel condition state can finally be determined. The channel is assumed to be Rayleigh: the SNR is exponentially distributed and the probability density function is as described below.

$$p(\gamma_i^j) = \frac{1}{\bar{\gamma}_i^j} \exp\left(-\frac{\gamma_i^j}{\bar{\gamma}_i^j}\right), \quad \gamma_i^j \geq 0, \tag{4.7}$$

where γ_i^j is the received SNR of the i-th SU on the j-th channel, and $\bar{\gamma}_i^j$ is the average SNR. The steady state probability of the k-th state of the i-th SU on the j-th channel can be calculated as given below:

$$\pi^j_{\text{CS,i}}(k) = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma_i^j)\,\mathrm{d}x. \tag{4.8}$$

Let $\boldsymbol{P}^j_{\text{CSi}} = \begin{pmatrix} p_{0,0} & \cdots & p_{0,N^j_{\text{SNR}}-1} \\ \vdots & \ddots & \vdots \\ p_{N^j_{\text{SNR}}-1,0} & \cdots & p_{N^j_{\text{SNR}}-1,N^j_{\text{SNR}}-1} \end{pmatrix}$ denote the state transition probability matrix of the

i-th SU on the j-th channel, where $p_{a,b}$ denotes the transition probability from state a to state b. According to [67], the transition probabilities can be obtained by:

$$p_{k,k+1} \approx \frac{f_m T_p}{\pi^j_{\mathrm{CS,i}}(k)} \sqrt{\frac{2\pi\Gamma_{k+1}}{\bar{\gamma}^j_i}} \exp\left(-\frac{\Gamma_{k+1}}{\bar{\gamma}^j_i}\right), \tag{4.9}$$

$$p_{k,k-1} \approx \frac{f_m T_p}{\pi^j_{\mathrm{CS,i}}(k)} \sqrt{\frac{2\pi\Gamma_{k}}{\bar{\gamma}^j_i}} \exp\left(-\frac{\Gamma_{k}}{\bar{\gamma}^j_i}\right), \tag{4.10}$$

where f_m is the maximum Doppler frequency caused by the movement of the SU mobile station. Here, The transition is assumed to occur only between adjacent states: $p_{j,k} = 0$, for any $|j-k| > 1$, and $p_{k,k} = 1 - p_{k,k+1} - p_{k,k-1}$. Then, using (5.3)-(5.7), the transition probability matrix $\boldsymbol{P}^j_{\mathrm{CS}}$ can be obtained. Let $V_{\mathrm{rate}}(k)$ represent the number of packets that can be transmitted in one time slot when the channel condition state is k. According to the rate in Table 4.1, $V_{\mathrm{rate}}(k), k \in \{0,1,2,3\} = \{0,1,2,3\}$.

As the error recovery method, a stop-and-wait ARQ is applied. The average packet error rate is Per. After data transmission, all packets arriving in the duration of the time slot will enter the SU buffer. If the total number of packets exceeds the buffer size, the overflow packets will be rejected.

4.3.4 Queueing model setup

4.3.4.1 Arrival process

The number of arrival packets at each SU during one time slot is assumed to follow a Batch Bernoulli process [46]. The probability that j packets arrive within one time slot at the i-th SU is denoted by $\alpha_i(j)$. Hence, a probability vector $\boldsymbol{\alpha}_i$ to describe the process can be written as below:

$$\boldsymbol{\alpha}_i = \big(\alpha_i(0), \alpha_i(1), \cdots, \alpha_i(v_i)\big), \tag{4.11}$$

where v_i is the maximum number of packets that are able to arrive at the i-th SU in one time slot and $\sum_{j=0}^{v_i} \alpha_i(j) = 1$.

4.3.4.2 State space

To establish the queueing model of the system, the number of packets in the buffer of each SU and the channel states are used to described the state of the system. The number of packets in the buffer of each SU is denoted by:

$$\boldsymbol{b} \in \mathcal{B} = \big\{(b_1, b_2, \cdots, b_M)|b_i \in \{0, 1, \cdots, K_i\}, i \in [1, M]\big\}, \tag{4.12}$$

where b_i denotes the number of packets in the buffer of the i-th SU. The channel state is composed of all J channel condition states to all M SUs and the PU occupancy state of all J channels. Here, the SNR of every PU channel is assumed to be identical to one SU. Thus, the channel condition state of J channels to M SUs is denoted by:

$$\boldsymbol{c} \in \mathcal{C} = \big\{(c_1, c_2, \cdots, c_M)|c_i \in [1, N_{\text{SNR}}(i)]\big\}, \tag{4.13}$$

where c_i represents the channel condition states of the PU channels, and $N_{\text{SNR}}(i)$ represents the number of channel condition states of the i-th SU. Thus, together with the PU occupancy state in Equation (4.4), the channel state of the system can be described by $\{\boldsymbol{o}, \boldsymbol{c}\}$, and the number of the channel states is $N_{\text{CS}} = 2^J \cdot \prod_{i=1}^{M} N_{\text{SNR},i}$.

The state space of the system can be denoted as below:

$$\Phi \triangleq \{(\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c})|\boldsymbol{b} \in \mathcal{B}, \boldsymbol{o} \in \mathcal{O}, \boldsymbol{c} \in \mathcal{C}\}. \tag{4.14}$$

The transition matrix between all the channel states $\{\boldsymbol{o}, \boldsymbol{c}\}$ is denoted by a matrix: $\boldsymbol{B}$ and $\boldsymbol{B}$ can be obtained as below:

$$\boldsymbol{B} = (\boldsymbol{P}_{\text{PU}}^1 \otimes \cdots \otimes \boldsymbol{P}_{\text{PU}}^J) \otimes \boldsymbol{P}_{\text{CS},1} \otimes \cdots \otimes \boldsymbol{P}_{\text{CS},M}, \tag{4.15}$$

where $\boldsymbol{P}_{\text{CS},i}$ is the channel condition state transition probability matrix of the i-th SU in Section 4.3.3.3 and $\otimes$ is the Kronecker product.

4.3.4.3 Transition Matrix

To derive the transition matrix, the distribution of the number of packets that can be transmitted needs to be derived first. According to the applied AMC scheme, the distribution of the number of transmitted packets of each SU through each channel can be obtained given the channel condition state $\boldsymbol{c}$. The number of transmitted attempts of the i-th SU through the j-th channel is denoted as C_i^j. According to the applied ARQ settings, the probability that x packets are transmitted successfully can be calculated as follows:

$$\mu(x) = \begin{cases} \binom{C_i^j}{x}(1-\mathrm{Per})^x & \text{for } 0 \le x \le C_i^j \\ \mathrm{Per}^{C_i^j} & \text{for } x = 0 \end{cases}, \tag{4.16}$$

where $\binom{C_i^j}{x}$ is C_i^j choose x. Given the allocation results: $\boldsymbol{A} = (a_1, \cdots, a_J)$, the distribution of the number of packets that can be transmitted by the i-th SU through the j-th channel can be obtained:

$$\boldsymbol{\mu}_i^j = \begin{cases} \left(\mu_i^j(0), \ldots, \mu_i^j(C_i^j)\right) & \text{for } a_j = i \\ (0) & \text{otherwise} \end{cases}. \tag{4.17}$$

Then the distribution of the number of packets that can be transmitted by the i-th SU given the channel state $\{\boldsymbol{o}, \boldsymbol{c}\}$ can be obtained as follows:

$$\boldsymbol{\mu}_i(\boldsymbol{A}, \boldsymbol{o}, \boldsymbol{c}) = \boldsymbol{\mu}_i^1 * \boldsymbol{\mu}_i^2 * \cdots * \boldsymbol{\mu}_i^J, \tag{4.18}$$

where $*$ is the convolution.

Then, the transition matrix that describes that the number of packets in the i-th SU's buffer changes from b_i to b_i' under $\boldsymbol{A}$ can be derived. First, a $(1+V) \times (1+C)$ auxiliary matrix that can select every possible arrival process and service process making every possible i-th SU's buffer changes is constructed. Here, V is the maximum number of arrival packets and C is the maximum number of transmitted packets during one time slot. The $(1+v, 1+c)$-th element of the matrix is derived as follows:

$$\Gamma_{v,c}(i, i') = \begin{cases} 1 & \text{when } i' = \min(v + i - \min(i, c), K) \\ 0 & \text{otherwise} \end{cases}. \tag{4.19}$$

Thus, the transition matrix that describes every possible changes of the i-th SU's buffer can be written as follows:

$$\boldsymbol{P}(b_i, b_i', \boldsymbol{A}) = \begin{bmatrix} \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(0,0)\boldsymbol{\mu}_i(\boldsymbol{A}) \cdots \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(0,K)\boldsymbol{\mu}_i(\boldsymbol{A}) \\ \vdots \quad \ddots \quad \vdots \\ \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(K,0)\boldsymbol{\mu}_i(\boldsymbol{A}) \cdots \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(K,K)\boldsymbol{\mu}_i(\boldsymbol{A}) \end{bmatrix}, \tag{4.20}$$

Let vector $\boldsymbol{b} = (b_1, b_2, \cdots, b_M)$ denote the number of packets in each SU's buffer at the current time slot, and $\boldsymbol{b}' = (b'_1, b'_2, \cdots, b'_M)$ denote the number of packets at the next time slot. Then, a coefficient matrix given $\boldsymbol{b}$, $\boldsymbol{b}'$ and $\boldsymbol{A}$ can be obtained:

$$\boldsymbol{T}(\boldsymbol{b},\boldsymbol{b}',\boldsymbol{A}) = \boldsymbol{P}(b_1,b'_1,\boldsymbol{A}) \otimes \boldsymbol{P}(b_2,b'_2,\boldsymbol{A}) \otimes \cdots \otimes \boldsymbol{P}(b_M,b'_M,\boldsymbol{A}). \tag{4.21}$$

Using the channel allocation protocol described by the DPM mechanism (e.g. Algorithm 2), a matrix D can be obtained given a system state $\phi \in \Phi$, and using the elements in $\boldsymbol{D}$ the following can be obtained:

$$\boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\phi) = \sum_{\mathrm{A}\in\mathcal{A}} \left(\prod_{j=1}^{J} d^j_{a_j} \boldsymbol{T}(\boldsymbol{b},\boldsymbol{b}',\boldsymbol{A}) \right), \tag{4.22}$$

where d^j_i is from the DPM D and $\mathcal{A}$ is the set of all possible allocation results. All the coefficients according to the index of channel condition states are listed and extended into a matrix with the same dimensions as $\boldsymbol{B}$:

$$\boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(1) \\ \vdots \\ \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\phi) \\ \vdots \\ \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(N_{\mathrm{CS}}) \end{bmatrix} \otimes \mathbf{1}_{1\times(\prod_{i=1}^{M}(K_i+1)\cdot N_{\mathrm{CS}})}. \tag{4.23}$$

To enumerate all coefficient matrices, β_{index} is mapped to $\boldsymbol{b} = (b_1, b_2, \cdots, b_M)$ using the following equation:

$$\text{index} = \sum_{i=1}^{M} b_i \times \text{base}_{i-1}, \tag{4.24}$$

where $\text{base}_i = \prod_{j=i}^{M}(K_i + 1)$.

The transition matrix can be divided into $\prod_{i=1}^{M}(K_i+1) \times \prod_{i=1}^{M}(K_i+1)$ blocks and obtained as follows:

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{T}_{\beta_0,\beta_0} \circ \boldsymbol{B} \ldots \boldsymbol{T}_{\beta_0,\beta_B} \circ \boldsymbol{B} \\ \vdots \quad \ddots \quad \vdots \\ \boldsymbol{T}_{\beta_B,\beta_0} \circ \boldsymbol{B} \ldots \boldsymbol{T}_{\beta_B,\beta_B} \circ \boldsymbol{B} \end{bmatrix}, \tag{4.25}$$

where $\circ$ is the entrywise product.

4.3.5 Performance evaluations

After the transition matrix $\boldsymbol{P}$ has been obtained, the steady probability vector of system states can be derived. The steady probability vector is denoted as $\boldsymbol{\pi}'$ and can be obtained using the QR algorithm, given the transition matrix $\boldsymbol{P}$. $\boldsymbol{\pi}'$ is a 1 by $\prod_{i=1}^{M}(K_i+1)\times N_{\mathrm{CS}}$ vector. A vector $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ can be mapped to any system state according to the sequences of the states in $\boldsymbol{P}$. Thus, the steady probability vector can be reorganized in the form of $\pi(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$.

4.3.5.1 Queue Length Distribution

The distribution of the expected number of packets that are in the i-SU's buffer at the beginning of a time slot can be derived using $\pi(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Let $P(q_i=l)$ denote the probability that l packets are waiting to be transmitted in the i-th SU's buffer, and $P(q_i=l)$ can be calculated as indicated below:

$$P(q_i=l)=\sum_{\boldsymbol{b}\in B_i(l)}\boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}), \tag{4.26}$$

where $B_i(l)=\{(b_1,\cdots,b_M)|b_i=l,b_j\in[0,K_j],j\neq i\}$ represents the set of all possible states of SU buffers when the i-th SU has exactly l packets in the buffer. The average queue length of the i-th SU can then be calculated as:

$$\overline{q_i}=\sum_{l=1}^{K_i} l\cdot P(q_i=l). \tag{4.27}$$

4.3.5.2 Average Throughput

In this section, the average throughput of the SU is described by the expected number of packets that can be transmitted in one time slot. Firstly, the distribution of the number of transmissions that the i-th SU can make under each system state $\phi=(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})\in\Phi$ can be calculated, which is denoted by $\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ as follows:

$$\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})=\sum_{A\in\mathcal{A}^j(i)} d_i^j\boldsymbol{\mu}_i(A,\boldsymbol{o},\boldsymbol{c}), \tag{4.28}$$

where d_i^j is from the DPM $\boldsymbol{D}$ obtained from ϕ and $\mathcal{A}^j(i)$ is the set of allocation results when the j-th channel is allocated to the i-th SU. Then, the average throughput can be obtained under $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$:

$$\mathrm{Thr}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) = \sum_{\beta} \min(\beta, b_i) \cdot t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},\beta), \tag{4.29}$$

where $t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},\beta)$ is the $(\beta+1)$-th element of $\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$, which is the probability that β transmissions have been made by the i-th SU under $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Thus the average throughput can be derived as:

$$\overline{\mathrm{Thr}} = \sum_{\phi\in\Phi} \mathrm{Thr}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) \cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}). \tag{4.30}$$

4.3.5.3 Average rejection

The average rejection in this part is denoted by the expected number of packets that are rejected at the beginning of one time slot. Firstly, the average number of rejected packets can be calculated under $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$:

$$\begin{aligned}\mathrm{Rej}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) = \sum_{y=0}^{v_i}\sum_{\beta} &\max(0, b_i - \beta + y - K_i) \\ &\cdot \min(\beta, b_i) \cdot t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},\beta).\end{aligned} \tag{4.31}$$

Thus, the average number of rejected packets can be derived as:

$$\overline{\mathrm{Rej}} = \sum_{\phi\in\Phi} \mathrm{Rej}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) \cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}). \tag{4.32}$$

4.3.5.4 Average delay

The delay performance is described by the expected time to transmit all packets that are in the buffer at the beginning of a time slot. In order to calculate it, the state space of the system is extended with another state f. f represents the number of packets that have already been transmitted by a selected SU. The extended system state is $\Phi' \triangleq \{\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},f\}$. Then, modification of the transition matrix based on $\boldsymbol{P}$, given that there are F packets in the selected SU's buffer, should be made. An extending band matrix given $\{\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}\}$ is built and $\boldsymbol{t}_i$ obtained, as in Equation (4.28):

$$\boldsymbol{P}_{\mathrm{ex}}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},f) = \begin{bmatrix} P_{0,0} \cdots P_{0,f} \\ \vdots \ \ddots \ \vdots \\ P_{f,0} \cdots P_{f,f} \end{bmatrix}, \tag{4.33}$$

where

$$P_{j,j'} = \begin{cases} t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},j'-j) & \text{for } j'-j \in [0, \max(\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}))] \\ 0 & \text{otherwise} \end{cases},$$

where $\max(\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}))$ is the maximum number of packets that can be transmitted under $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Then extend $\boldsymbol{P}$ row by row as follows:

$$\boldsymbol{P}'_{\text{delay}}(f) = \begin{bmatrix} \boldsymbol{P}(\phi(0)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(0), f) \\ \vdots \\ \boldsymbol{P}(\phi(R_m)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(R_m), f) \end{bmatrix}, \tag{4.34}$$

where $R_m = \prod_{i=1}^{M}(K_i+1)\cdot N_{\text{CS}}$ is the size of $\boldsymbol{P}$, and $P(\phi(r))$ is the r-th row of matrix $\boldsymbol{P}$. The absorbing states are the states when all the packets in the buffer are transmitted ($b_i = 0$ or $f = F$). Then, all the absorbing states are removed from $\boldsymbol{P}'_{\text{delay}}$ and a matrix $\boldsymbol{P}_{\text{delay}}$ is got. The average first passage time to the absorbing states under each state is denoted as $\boldsymbol{S}$, which is a $(K_s \cdot \prod_{i\neq s}(K_i+1)\cdot N_{\text{CS}})$-by-1 column vector ($s$ is the label of the selected SU). Then solve $\boldsymbol{S}$ from the following equations:

$$(\boldsymbol{I} - \boldsymbol{P}_{\text{delay}}) \cdot \boldsymbol{S} = \mathbf{1} \tag{4.35}$$

where $\boldsymbol{I}$ is the identity matrix, $\mathbf{1}$ is a column vector whose elements are all 1 and size is the same as $\boldsymbol{S}$. Let $\boldsymbol{S}(F)$ be the solution of $\boldsymbol{S}$ given F. Thus the average first passage time from the state of F packets in the buffer to the empty buffer state is:

$$d(F) = \sum_{\boldsymbol{b}\in B_s(F)} S(F,\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) \cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}), \tag{4.36}$$

where $B_s(F)$ is the set of all states of SU buffers when the selected SU's buffer has F packets in it. Finally, the average delay of the i-th SU can be obtained as below:

$$\overline{\text{Dly}} = \sum_{t=1}^{K_i} d(t). \tag{4.37}$$

4.4 MAXIMUM THROUGHPUT PROTOCOL

4.4.1 Basic ideas

The proposed MT protocol not only focuses on the capacity of available PU channels to the SUs, but also takes the transmission demands of SU into consideration. The MT protocol allocates a channel to the SU that has the potential to transmit most packets in the following time slot. Here, the potential throughput of one SU is the minimum between the channel capacity according to the channel condition state and AMC (ignoring the effect of ARQ) and the transmission demands, which are represented by number of packets in the buffer. Moreover, in order to minimize the probability of the rejection of the arrival packets, when two or more SUs have the identical transmission potential, the MT protocol will allocate the channel to the SU whose buffer is more likely to be full afterwards. To measure the possibility, a block probability (BPI) is set up to describe how close the state of an SU's buffer is to the "full" state. Finally, the DPM mechanism is used to describe the allocation results.

4.4.2 Calculation of the potential throughput

In this M-SU N-channel CRN, a vector $A = (a_1, a_2, \cdots, a_J)$ is used to describe the allocation result, where $a_j = 0$ represents that the j-th channel is not allocated to the SU system and $a_j = i, i \in \{1, 2, \cdots, M\}$ represents that the j-th channel is allocated to the i-th SU. The capacity of the j-th channel to the i-th SU can be obtained given the channel condition state and the modulation schemes. This capacity is denoted by cap_i^j. Thus, the overall capacity of all the channels that are allocated to the i-th SU can be obtained, which is denoted by cap_i as follows.

$$\text{cap}_i = \sum_{j \in \{j | a_j = i\}} \text{cap}_i^j. \tag{4.38}$$

The potential throughput can be denoted as $\text{Thr}_i = \min(b_i, \text{cap}_i)$, where b_i is the number of packets in the i-th SU's buffer. Finally, the potential throughput of the given allocation result A can be obtained as:

$$\text{Thr}(A) = \sum_{i=1}^{M} \text{Thr}_i. \tag{4.39}$$

4.4.3 Block probability index

The BPI in the MT protocol is defined as the average first passage time from the current state to the full buffer state, ignoring the influence of the PU activity. To calculate the BPI, a $\text{Geo}^\text{X}/\text{Geo}^\text{Y}/1/\text{K}$ queueing model is applied. The algorithm takes the distribution of the number of arrival packets, the distribution of the number of packets that can be transmitted, the size of the buffer and the current number of packets in the buffer as the input. The detail of the calculation is shown in Algorithm 1.

Algorithm 1 Algorithm for block probability index index $\boldsymbol{I}_b$

Input:

Arrival distribution: $\boldsymbol{\alpha} = \{\alpha_0, \alpha_1, \cdots, \alpha_v\}$
Service distribution: $\boldsymbol{\beta} = \{\beta_0, \beta_1, \cdots, \beta_c\}$
Buffer size: K
Number of packets in buffer: b

Output: Block probability index: I_b

```
for i_1 = 0 to K − 1 do
  for i_2 = 0 to K − 1 do
    for i_3 = 0 to v do
      for i_4 = 0 to c do
        if i_2 = min(v + i_1 − min(i_1, c), K) then
          Let Γ_{i_3,i_4}(i_1, i_2) = 1
        else
          Let Γ_{i_3,i_4}(i_1, i_2) = 0
        end if
      end for
    end for
  end for
end for
```

Let matrix $\boldsymbol{\Gamma}(i,i') = \begin{pmatrix} \Gamma_{0,0}(i,i') \cdots \Gamma_{v,0}(i,i') \\ \vdots \quad \ddots \quad \vdots \\ \Gamma_{0,c}(i,i') \cdots \Gamma_{v,c}(i,i') \end{pmatrix}$

Let matrix $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{\alpha\Gamma}(0,0)\boldsymbol{\beta} & \cdots & \boldsymbol{\alpha\Gamma}(0,K-1)\boldsymbol{\beta} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\alpha\Gamma}(K-1,0)\boldsymbol{\beta} & \cdots & \boldsymbol{\alpha\Gamma}(K-1,K-1)\boldsymbol{\beta} \end{pmatrix}$

Let column vector $\boldsymbol{I}_b = \begin{pmatrix} I_b(0) \\ I_b(1) \\ \vdots \\ I_b(K-1) \end{pmatrix}$

Solve equation group: $(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{I}_b = \boldsymbol{1}$

$I_b = I_b(b)$.

4.4.4 Algorithm of the allocation protocol

From the PU occupancy state: $\boldsymbol{o}$, all possible channel allocation results can be listed. Thus, given the capacities of PU channels to the SU system ($\mathcal{C} = \begin{bmatrix} \mathrm{cap}_1^1 \cdots \mathrm{cap}_M^1 \\ \vdots \quad \ddots \quad \vdots \\ \mathrm{cap}_1^J \cdots \mathrm{cap}_M^J \end{bmatrix}$) and the number of packets in the SUs' buffers, the potential throughput of each allocation result can be calculated. Let $\mathcal{A}_{\max}$ denote the set of allocation results with the maximum potential throughput. Let L_i denote all the possible allocation results of the i-th channel in the set $\mathcal{A}_{\max}$. Then calculate the BPI of all SUs in the set

Algorithm 2 Algorithm for channel allocation matrix $\boldsymbol{D}$

Input:

Channel capacities to SU system: $\mathcal{C}$
Number of packets in SU buffer: $\mathcal{B}$
PU occupancy state: $\mathcal{O}$
Number of SUs: M
Number of PU channels: J

Output: Distribution probability matrix:

$$\boldsymbol{D} = \begin{pmatrix} d_0^1 \, d_1^1 \cdots d_M^1 \\ \vdots \;\; \vdots \;\; \ddots \;\; \vdots \\ d_0^J \, d_1^J \cdots d_M^J \end{pmatrix}$$

Let $\boldsymbol{a}_j, 0 \leq j \leq J$ be sets that represent the index of the SUs that can access channel j.
Set $d_i^j = 0$ for $i \in \{0, 1, \cdots, M\}, j \in \{1, 2, \cdots, J\}$
for $j = 1$ to J **do**
  if $O^j = 0$ **then**
    $\boldsymbol{a}_j = \{0\}$, $d_0^j = 1$
  else
    $\boldsymbol{a}_j = \{1, 2, \cdots, M\}$
  end if
end for
Let $\mathcal{A} = \{A = (a_1, a_2, \cdots, a_J) | a_j \in \boldsymbol{a}_j\}$
Calculate $\mathcal{T} = \{T = \mathrm{Thr}(A) \, | A \in \mathcal{A}\}$.
Find the maximum in $\mathcal{T}$: $T_{\max}$.
Let $\mathcal{A}_{\max} = \{A | \mathrm{Thr}(A) = T_{\max}\}$.
for $j = 1$ to J **do**
  Let L_j be the set of all values of a_j, $\forall A \in \mathcal{A}_{\max}$.
  Find the minimum in $\boldsymbol{I}^j = \{I = I_b(l) | l \in L_j\}$: $I_{\min}^j$
  Let $\boldsymbol{d}^j = \{d | I_b(d) = I_{\min}^j\}$. N_d^j is the number of members in $\boldsymbol{d}^j$.
  for all $i \in \boldsymbol{d}^j$ **do**
    Let $d_i^j = \frac{1}{N_d^j}$
  end for
end for

L_i. The distribution of the number of packets that each SU can transmit is calculated as described below. Assume that all the PU channels are utilized by the SU under the current SNR. Then using the probability of the channel condition state and the AMC transmission rate in Section 4.3.3.3, the distribution can be obtained. Finally, the SUs in the set L_i that have the minimum BPI share the opportunity to access the i-th channel. Details are shown in Algorithm 2.

Table 4.2. Comparison of the runtime complexity of different channel allocation algorithms.

Protocols	Best case	Worst case
MT	$O(J\cdot \max(M^2, K^2vc, K^4))$	$O(M\cdot J\cdot \max(M^J, K^2vc, K^4))$
MR	$O(M^2\cdot J)$	$O(M^{J+1}\cdot J)$
MT no Index	$O(M^2\cdot J)$	$O(M^{J+1}\cdot J)$

4.4.5 Complexity analysis of the algorithm

The calculation of the potential throughput and the BPI in the proposed MT algorithm increases the runtime complexity of the algorithm compared to the algorithms that ignore them. For comparison, a MR protocol and a modified MT protocol (MT no index) in which the BPI is removed are introduced. The details are given below.

4.4.5.1 The maximum rate protocol

The MR protocol is based on the principle that a PU channel is allocated to those SUs having the maximum transmission rate through that channel. The calculation of function $\text{Thr}_{\max}$ is similar to the potential throughput in Equation (4.39). Given the channel condition state and the modulation schemes, the total capacity of all the channels that are allocated to the i-th SU can be obtained, which is denoted by cap_i. The throughput of the given allocation result A under MR is calculated as:

$$\text{Thr}_{\text{MR}}(A) = \sum_{i=1}^{M} \text{cap}_i. \tag{4.40}$$

The details of the algorithm are listed in Algorithm 3.

4.4.5.2 The MT without BPI

To provide a reference for comparison, a modified version of the MT protocol in which the BPI part is removed is proposed. This modified protocol can help provide a reference to investigate the trade-offs between the performance and the computational complexity during the channel allocation algorithm design. The details of the algorithm are listed in Algorithm 4.

Algorithm 3 Algorithm for channel allocation matrix $\boldsymbol{D}$ of MR

Input: Channel capacities to SU system: $\mathcal{C}$
Number of packets in SU buffer: $\mathcal{B}$
PU occupancy state: $\boldsymbol{o}$
Number of SUs: M
Number of PU channels: J

Output: Distribution probability matrix:

$$\boldsymbol{D} = \begin{pmatrix} d_0^1 \, d_1^1 \cdots d_M^1 \\ \vdots \;\; \vdots \;\; \ddots \;\; \vdots \\ d_0^J \, d_1^J \cdots d_M^J \end{pmatrix}$$

Let $\boldsymbol{a}_j, 0 \leq j \leq J$ be sets that represents the index of SU that can access channel j.
Set $d_i^j = 0$ for $i \in \{0, 1, \cdots, M\}, j \in \{1, 2, \cdots, J\}$
for $j = 1$ to J **do**
  if $O^j =$ busy **then**
    $\boldsymbol{a}_j = \{0\}$, $d_0^j = 1$
  else
    $\boldsymbol{a}_j = \{1, 2, \cdots, J\}$
  end if
end for
Let $\mathcal{A} = \{A = (a_1, a_2, \cdots, a_J) | a_j \in \boldsymbol{a}_j\}$
Calculate $\mathcal{T} = \{T = \mathrm{Thr}_{\mathrm{MR}}(A) \, | A \in \mathcal{A}\}$.
Find the maximum in $\mathcal{T}$: $T_{\max}$.
Let $\mathcal{A}_{\max} = \{A | \mathrm{Thr}_{\mathrm{MR}}(A) = T_{\max}\}$.
for $j = 1$ to J **do**
  Let L_j be the set of all values of a_j, $\forall A \in \mathcal{A}_{\max}$.
  N_d^j is the number of members in L_j.
  for all $i \in L_j$ **do**
    Let $d_i^j = \frac{1}{N_d^j}$
  end for
end for

4.4.5.3 Comparison of the runtime complexity

In this section, the complexity of the protocols mentioned in this chapter are compared. The details are listed in Table 4.2. From the complexity comparison, it can be found that the contribution of the potential throughput and the BPI to the complexity is related to the following parameters: the SU buffer size K, the maximum arrivals v and the number of channel states c. The effect of the parameters is significant when the number of SUs (M) and the number of PU channels (J) are small. M and J are the determining factors of the complexity of all algorithms when they become large. The complexity of the algorithms increases when the number of available PU channels increases. The best case in Table 4.2 is that there is only one PU channel available to the SU, while the worst case is that all J channels are available. The numerical results of the runtime statistics and performance comparison are

Algorithm 4 Algorithm for channel allocation matrix $\boldsymbol{D}$ of MT without BPI

Input: Channel capacities to SU system: $\mathcal{C}$
Number of packets in SU buffer: $\mathcal{B}$
PU occupancy state: $\boldsymbol{o}$
Number of SUs: M
Number of PU channels: J

Output: Distribution probability matrix:

$$\boldsymbol{D} = \begin{pmatrix} d_0^1 \, d_1^1 \cdots d_M^1 \\ \vdots \ \ \vdots \ \ddots \ \vdots \\ d_0^J \, d_1^J \cdots d_M^J \end{pmatrix}$$

Let $\boldsymbol{a}_j, 0 \leq j \leq J$ be sets that represents the index of SU that can access channel j.
Set $d_i^j = 0$ for $i \in \{0, 1, \cdots, M\}, j \in \{1, 2, \cdots, J\}$
for $j = 1$ to J **do**
  if $O^j =$ busy **then**
    $\boldsymbol{a}_j = \{0\}$, $d_0^j = 1$
  else
    $\boldsymbol{a}_j = \{1, 2, \cdots, J\}$
  end if
end for
Let $\mathcal{A} = \{A = (a_1, a_2, \cdots, a_J) | a_j \in \boldsymbol{a}_j\}$
Calculate $\mathcal{T} = \{T = \mathrm{Thr}(A) | A \in \mathcal{A}\}$.
Find the maximum in $\mathcal{T}$: $T_{\max}$.
Let $\mathcal{A}_{\max} = \{A | \mathrm{Thr}(A) = T_{\max}\}$.
for $j = 1$ to J **do**
  Let L_j be the set of all values of a_j, $\forall A \in \mathcal{A}_{\max}$.
  N_d^j is the number of members in L^j.
  for all $i \in L^j$ **do**
    Let $d_i^j = \frac{1}{N_d^j}$
  end for
end for

further shown and discussed in Section 4.5.

4.5 NUMERICAL RESULTS AND DISCUSSIONS

In this part, the MT protocol and other related protocols mentioned in the previous section are implemented. To reduce the complexity without loss of generality, a model with two users and two channels to show the performance of the MT protocol and the advantages of the analytical framework are established and analyzed. The default settings of this model are as follows:

- Channel condition: $P_{target} = 0.01$, $f_m = 10\,\mathrm{Hz}$
- PU activity model: $t_{PU} = 1$, $\theta = 50\%$

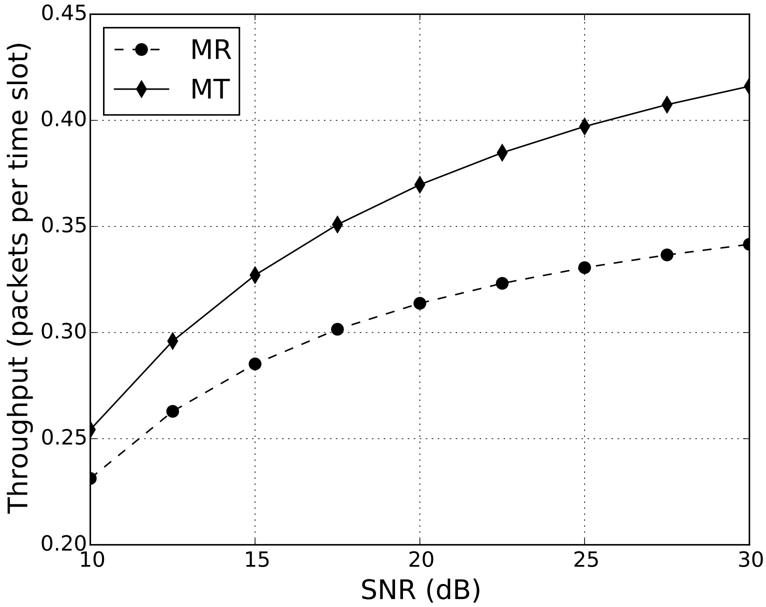


Figure 4.3. Throughput versus SNR under MT protocol and MR protocol

- SU buffer size: $K_1 = 6$, $K_2 = 6$
- Arrival process: $\alpha_1 = \alpha_2 = \{0.5, 0.5\}$
- ARQ setting: Per $= 0.01$.

4.5.1 Performance comparison with other protocols

Firstly, the numerical result that compares the throughput of the MT protocol and the MR protocol under different SNRs are shown in Figure 4.3. Applying the MT protocol to allocate the channels, 10% to 20% improvement on throughput can be achieved. The requirement of buffer size can also be improved by the application of the MT protocol. Figure 4.4 shows the queue length distributions under MT and MR protocols when the SNRs vary. Then, the PU activity parameters and SU arrival process parameters are set to multiple values to investigate how these parameters affect the improvement of the throughput. The results are shown in Figure 4.5. In Figure 4.5(a), θ is the PU occupancy probability of both channels; in Figure 4.5(b), α is the probability that no packets arrived during the previous time slot. The results show that the MT protocol will achieve greater improvement when the PU activity is more frequent or there are more SU arrival packets. It can be concluded through the results that the MT protocol is capable of outperforming the MR protocol on maximizing the throughput and more

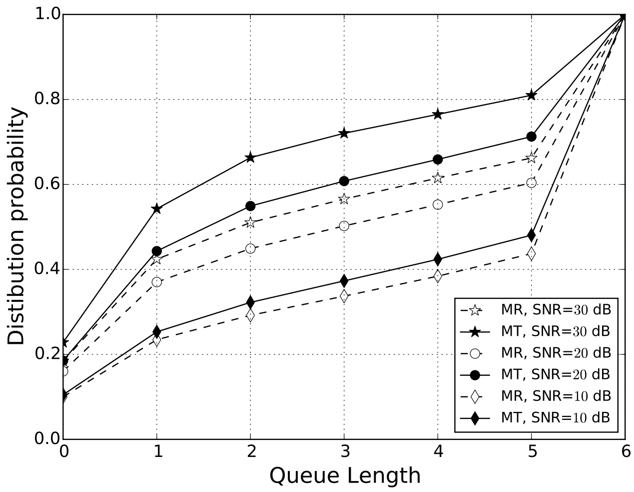


Figure 4.4. Queue length distribution under MT protocol and MR protocol

improvement can be achieved when the SU system has more available spectrum resources or the data traffic of the SU system becomes heavier.

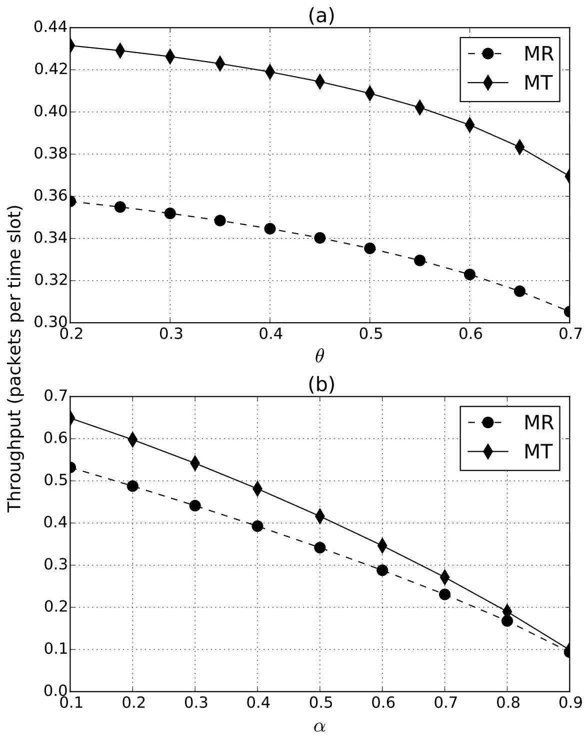


Figure 4.5. Throughput improvements versus different PU activity (θ) and arrival process (α) settings.

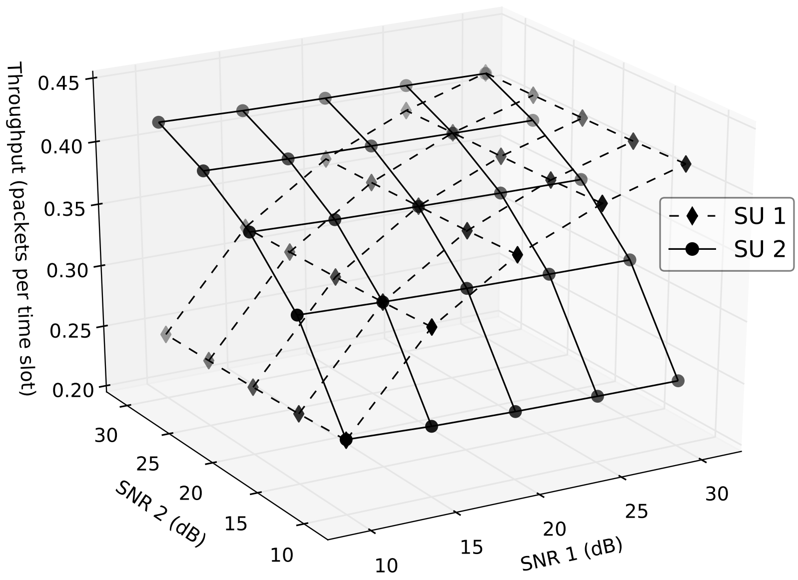


Figure 4.6. Throughput of SU 1 and SU 2 versus SNRs.

4.5.2 Performance analysis of multiple SUs

One advantage of this analytical framework is that all PU channels can be configured independently, and the performance of all SUs can be obtained simultaneously and individually. The influence of the channel allocation protocol on both SUs when the system parameters vary is investigated. Figure 4.6 depicts the throughput of SU 1 and SU 2 when the SNRs of the two SUs vary. The performance of the MT and MR protocols in terms of the throughput of the two SUs are compared under an selected SNR setting. The results are shown in Figure 4.7. One shortcoming of the MR protocol can be observed. When one of the SUs has a better transmission condition (higher SNR), the MR protocol or the protocols using similar ideas inclines to allocate the channels to this SU even if this SU does not have enough transmission demands (packets waiting in the buffer). This strategy will obviously waste precious spectrum resources. Moreover, the increase in the SNR of the SU under better transmission conditions will affect the performance of the SU under worse conditions. By applying the MT protocol, this influence can be eliminated. In Figure 4.8, the average delay of SU 1 and SU 2 under different SNRs is shown. The delay performance of the SU with better SNR remains the same while the other SU's SNR increases. THe results discussed in this section prove that using the MT protocol, the increase in one SU's SNR does not affect another SU's transmission much when the spectrum resource

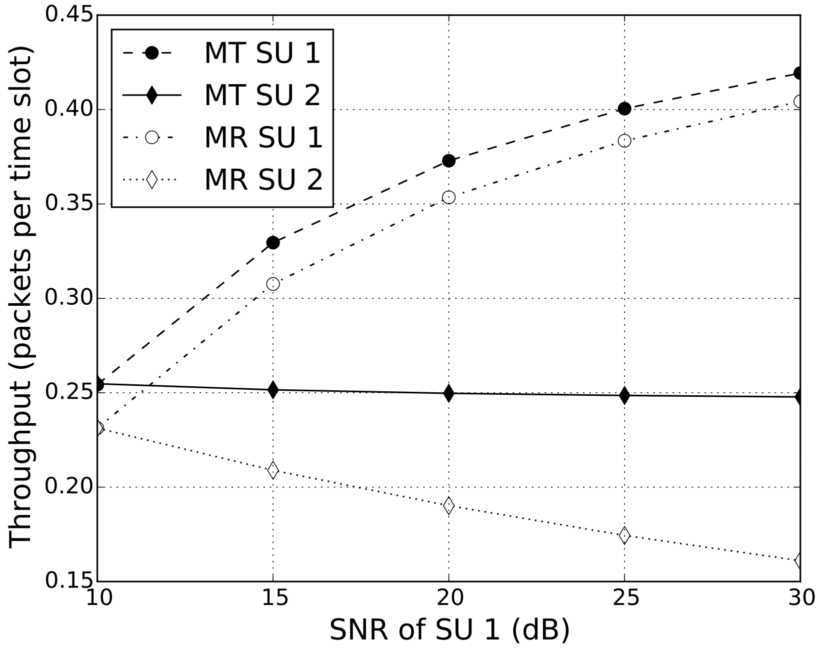


Figure 4.7. Throughput of two SUs versus SNR of SU 1. The SNR of SU 2 is 10 dB.

is sufficient. This also proves that the MT protocol is able to guarantee the transmission quality of each SU in the system when the spectrum resource is sufficient, thus achieve higher resource utilization efficiency besides improving throughput.

The following advantages of the framework are revealed by the results above. First, the proposed queueing analytical framework is able to model various CR channel allocation protocols and compare their performance in various situations. Second, applying this framework, the individual performances of each SU can be obtained simultaneously and efficiently. Moreover from the performance of each individual SU, it is possible to investigate how a channel allocation protocol works in more detail.

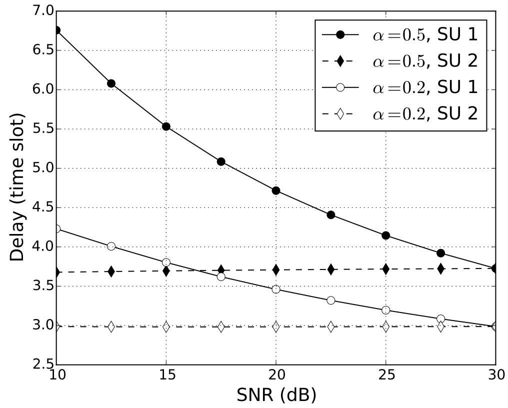


Figure 4.8. Average delay of SU versus SNR. The SNR of SU 2 is 30 dB. α is the probability that no packets arrive at SU 1 during the previous time slot.

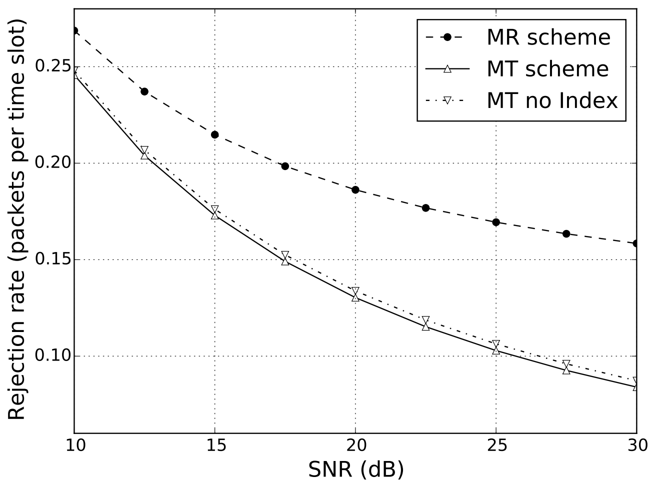


Figure 4.9. Packet rejection rate versus SNR under different protocols

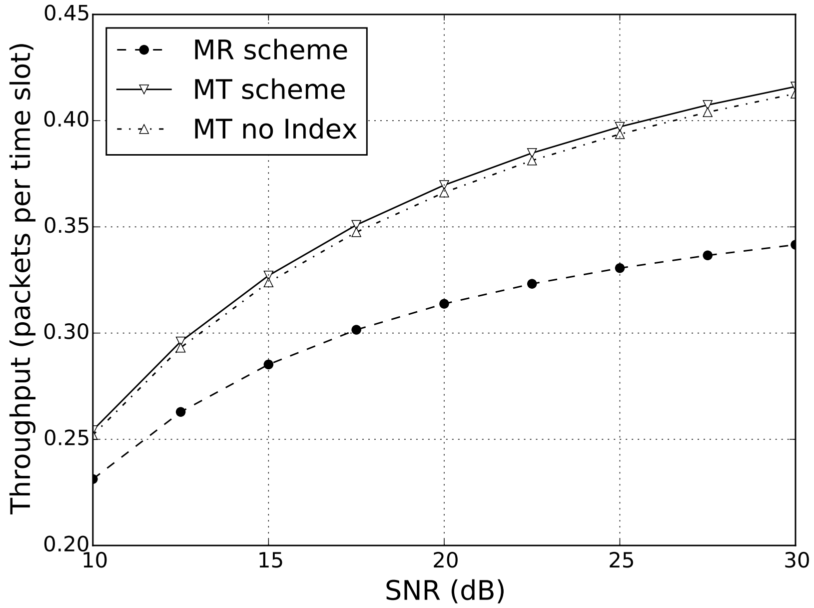


Figure 4.10. Throughput versus SNR under different protocols.

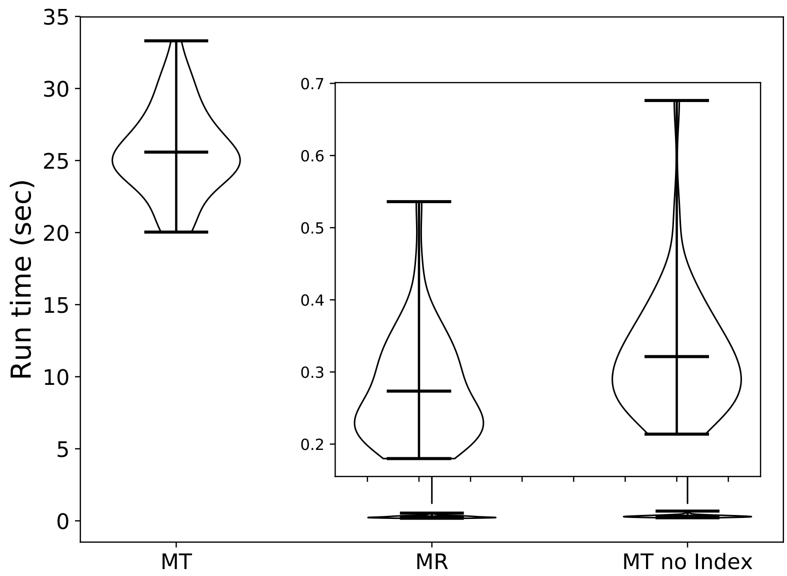


Figure 4.11. Run time statistics of 1000 times of simulations under different channel allocation protocols (violin plot).

4.5.3 Discussion on the trade-off between the protocols

In the process of designing a channel allocation protocol for the CRN, the trade-off between the performance and the complexity of the protocol is evident. The comparison of the MT protocol versus the MR protocol in this chapter can serve as an example. The calculation of the potential throughput and the BPI part of the MT protocol are the main reasons why it is able to outperform the MR protocol. However, the calculation also consumes a lot of computing resources of the SU system. Using the framework, the trade-off from some specific angles is revealed. The modified version of the MT protocol in Section 4.4.5.2 is applied as a reference. Because BPI is designed to improve the packet rejection rate, the throughput and the rejection rate under different SNR settings in Figure 4.9 and Figure 4.10 are compared. The three protocols are run in MATLAB 1000 times. The run time statistics are shown in Figure 4.11 in the form of violin plot. The results show that great computing resource consumption is caused by the application of BPI in the protocol. However, it does not yield much improvement. Thus, applying BPI is probably not a good option for the channel allocation protocol if the computing resources of the SU system are limited. Generally speaking, MT protocol consumes more computing resources than the MR protocol to achieve more throughput and better utilization of radio resource, which means the MT protocol is appropriate for delay sensitive networks. On the other hand, in delay tolerant applications, MR would work well if the computing resources is limited.

This part of the analysis proves that the proposed framework can serve as reference during the design and evaluation of the CR channel allocation protocols.

4.6 CONCLUSION

In this chapter, the DPM mechanism is applied to describe channel allocations with preset objectives. Based on the DPM, an improved channel allocation protocol, named the MT protocol is introduced. Compared to other channel allocation protocols with similar objectives, such as MR, it is able to improve the SU system performance of throughput by 10% to 20% under the selected system settings. Performance measures such as the queue length in the buffer and the packet rejection rate are also improved. The negative effects of SUs under poor transmission conditions in the system on the performance of one SU, such as throughput and packet delay, can also be reduced by the MT protocol.

To carry out the analysis of different channel allocation protocols, the author established a queueing model based on the DPM mechanism and it proves to be practicable to describe channel allocation protocols in queueing analysis frameworks. This framework is able to describe the behavior of each individual SU in the system independently and to evaluate various performance measures of each SU under various system and environment settings. This makes the framework capable of comparing different channel allocation protocols from more angles and in more detail.

The queueing analytical framework and the MT protocol in this chapter show that the DPM mechanism is able to describe the CR channel allocation protocols with preset objectives easily and to analyze them in an efficient and comprehensive way. As evidenced by this chapter, the DPM mechanism is able to play an important role in CR channel allocation protocol design and analysis.

CHAPTER 5 OPTIMIZATION OF CHANNEL ALLOCATION BASED ON DPM

5.1 CHAPTER OVERVIEW

To take the best advantage of the utilization efficiency of the radio resource provided by CR technology, especially in a complicated context such as a multi-user multi-channel CRN cell, a quantitative analytical model of the channel allocation protocol is a key issue. Since the access of the CR system to the PU channels is dynamic, the more efficiently the evaluations of the performance determining the channel allocation can be done, the better utilization of the radio resource can be achieved.

In this chapter, the DPM mechanism is used as channel allocation protocol that is able to deploy the radio resource quantitatively. Through the DPM, the best method to allocate the resources to achieve the optimization objective under various system settings in CRN can be investigated. As a practical case, an optimization goal is set up: to maximize an overall performance evaluation of the CR system using the concept of weighted throughput under delay constraints. The weight represents the importance of the CR user and is the feature that distinguishes different types of users. Then a queueing analytical framework based on DPM is built to model the CR system and comprehensive performance evaluations of every individual SU are obtained efficiently. A parameter named overtime probability (OP) is subsequently introduced to describe the measure of delay approximately, with high computing efficiency. The author thus formulates the optimization problem in CRN to maximize the overall weighted throughput under delay constraints represented by OP and an algorithm is developed to find the solution in terms of DPM. The numerical results reflected below reveal the optimal allocation to achieve the optimization objectives under various system settings and verify the computing efficiency of the proposed framework.

5.2 BACKGROUND

CR technology provides a solution to accommodate the rapid growth in the demand for wireless data transmission by utilizing the precious radio resources more efficiently. However, the way to evaluate improvement of radio resource utilization based on CR technology has not been studied well yet. Two preliminary tasks need to be accomplished to complete an evaluation of CR technology. First, because the CRN may carry different types of services or present users with various priorities compared to the traditional telecommunication system, a specialized performance evaluation criterion needs to be designed, considering the characteristics of the CRN. Second, because the access of a CR system to radio resources is more complicated than the traditional telecommunication system, an analytical framework is needed to model the protocol that the CRN uses to manage the radio resource that is accessible to it and make it possible to investigate the relationship between the radio resource management protocols and the performance of the CRN.

Whether the CR system can access the radio resources depends on the behavior of the PUs, thus the resources available to the CR system are limited and dynamic. This requires the CR system to have a criterion to evaluate the priorities among the SUs when the resource is insufficient to serve all the SUs. Among the research outputs that investigate how the importance of users should be determined, weighted throughput is appropriate for the CR. The weights admit various interpretations, including levels of importance, "utility" and price [122].

Using the DPM mechanism, a configurable and flexible channel allocation protocol can be generated to adapt to various situations. In addition, when dealing with multiple users in CRN, SUs are usually described by one identical model, which makes it complicated to analyze multiple SUs separately, especially when the SUs are of different kinds. To improve this, the authors developed a queueing analytical framework using the DPM, which is capable of modeling the CR system with every individual SU independently configurable. Therefore the performance of every individual SU can be obtained efficiently. Among the performance evaluations, the calculation of the measure of delay is complicated. Much research on delay analysis focuses on the process of transmission of the packets ahead of the tagged one, such as in [120]. This method is accurate; however, it requires consideration of every step of the transmission of the packets beforehand, which makes the computing time-consuming. Since the radio resources that a CR system can use are dynamic and performance evaluations need to be carried out to direct the allocation of the channels, to make the best use of the limited resources available to the CR system not only the accuracy of the performance evaluation is of concern, but also its efficiency. Considering this, the author introduces a method to estimate the measure of delay with high efficiency. Applying the proposed analytical framework, the relationship between the performance and channel allocation protocol configurations can be determined. Moreover, a method to achieve the optimal channel allocation to meet the requirements of performance is investigated. An optimization problem considering both the overall performance in terms of weighted throughput and the delay constrains that are used to guarantee the transmission of SUs with low priority is formulated to find the optimal channel allocation solutions. Analysis of the effect of the system settings and environmental parameters on the optimal allocation results is carried out afterwards.

5.3 OBJECTIVE FUNCTION FORMULATION

From the DPM mechanism, a group of channel allocation protocols can be generated and characterized by the a probability matrix. In order to optimize the performance of a CR system using DPM as channel allocation protocol, the function that relates the DPM settings to the performance metrics must be obtained. In the next part, the performance metrics that are potentially related to objectives of CRN are derived with the DPM settings as input under the queueing analytical framework proposed by the author.

5.3.1 System model and assumptions

5.3.1.1 Basic model settings

In this section, an overlay CRN with multiple SUs carrying different types of services inside a CRN cell or cluster is considered. An example of the system model is shown in Figure 5.1. Multiple PU channels are available to this SU system while there is no PU transmitting through them. Perfect control information exchange between the PU and SU system is assumed. The SU system comprises a base station (or a center node) and multiple SUs of different types working within the coverage of one or more PU networks. The number of SUs is denoted by M. The number of PU channels that are available to the SU system is denoted by J. Both the PU and SU systems work under the same time slot structure synchronized with each other. In this section, only the uplink data transmission is

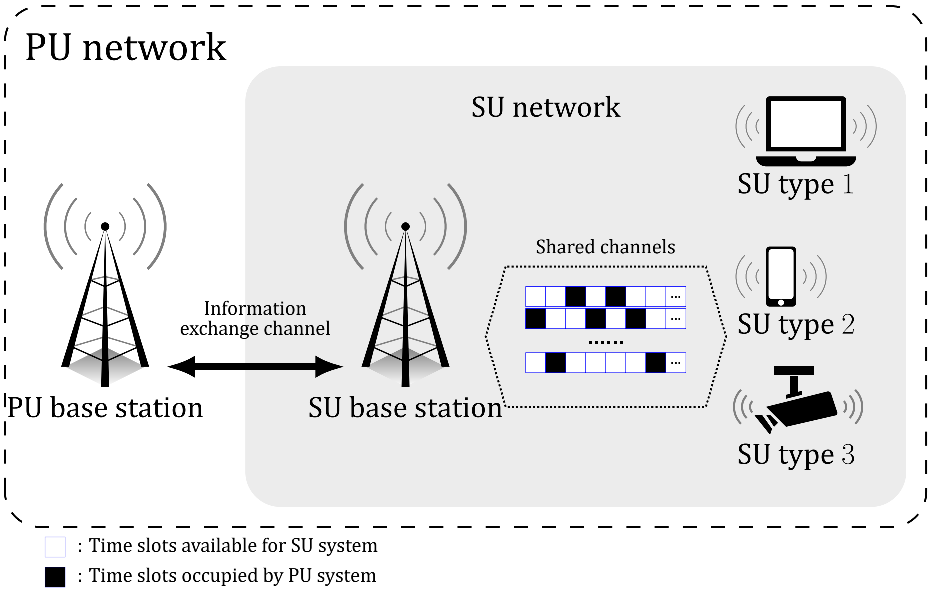


Figure 5.1. System model.

considered. Each SU has a finite buffer to store the packets waiting to be transmitted. The buffer size of the i-th SU is denoted by K_i.

The channel allocation protocol is the rule that the SU system uses to determine which SU can transmit on the channel where the PU is sensed to be absent. The common objective of the channel allocation protocol in the CRN is to fulfill QoS requirements such as maximizing the throughput, minimizing the packet transmission failure rate and minimizing the delay.

5.3.2 Primary user activity model

In an overlay CRN, the SU system can use a channel when there is a PU transmitting on it. In this research, the PU activities on a channel are modeled as a time-homogeneous first-order Markov process with two states. This process is independent of the behavior of the SU system. Let $O^j(t) \in \{0,1\}$ represent the PU activity state of the j-th channel at the t-th time slot. When there is a PU transmitting, the state is called "busy" and $O^j(t) = 0$, otherwise the state is called "free" and $O^j(t) = 1$. The transition matrix of the process can be written as:

$$\boldsymbol{P}_{\mathrm{PU}}^{j} = \begin{pmatrix} p_{\mathrm{b}\to\mathrm{b}}^{j} & 1 - p_{\mathrm{b}\to\mathrm{b}}^{j} \\ 1 - p_{\mathrm{f}\to\mathrm{f}}^{j} & p_{\mathrm{f}\to\mathrm{f}}^{j} \end{pmatrix}, \tag{5.1}$$

where $p^j_{\mathrm{b}\to\mathrm{b}} = \Pr\{O^j(t) = 0|O^j(t-1) = 0\}$ denotes the probability that there is a PU transmitting on the j-th channel (the "busy" state), given that there was PU transmitting on the channel (the "busy" state) at the previous time slot, and $p^j_{\mathrm{f}\to\mathrm{f}} = \Pr\{O^j(t) = 1|O^j(t-1) = 1\}$ denotes the probability that the j-th channel state becomes "free" from a "free" state.

5.3.3 Secondary system transmission model

The SU system transmits the data within time slots that are synchronized with the PU system. An SU time slot consists of three parts: the spectrum sensing part, the channel allocation part and the data transmission part.

The SU base station and the mobile stations keep sensing the existence of the PU transmission and the base station collects all the information at the beginning of a time slot to estimate whether the PU is absent on every PU channel. In the model, perfect spectrum sensing is assumed. Besides the PU's presence, the transmission conditions (e.g. the SNR at transceivers) of each SU mobile station are also collected by the SU base station and used to guide the channel allocation.

According to the estimation of available PU channels in the spectrum sensing part, the SU system allocates the available channels to the SUs to transmit data during the current time slot. The DPM protocol works as follows: a J-by-M DPM is :

$$\boldsymbol{D} = \begin{pmatrix} d_1^1 \cdots d_M^1 \\ \vdots \ \ddots \ \vdots \\ d_1^J \cdots d_M^J \end{pmatrix}, \tag{5.2}$$

where d_i^j represents the probability that the SU system allocates the j-th channel to the i-th SU when the channel is not occupied by the PU. Here, it is assumed that one PU channel can only be assigned to one SU, thus in $\boldsymbol{D}$, $\sum_{i=1}^{M} d_i^j = 1$ holds for any $j \in [1, J]$.

If a PU channel is allocated to one SU, the SU transmits its data during the rest of the time slot. To make better use of the channel resources, an AMC scheme is applied by the SU system. The modulation schemes are determined by the channel condition and the channel condition is modeled as a Markov process. A transition matrix of channel condition states $\boldsymbol{P}^j_{\mathrm{CS}}$ is built to represent this model. The details of the formulation of the matrix are given as below. The transmission mode on the physical layer is convolutionally coded M_n-ary rectangular or square QAM modes, which are adopted from the HIPERLAN/2 or IEEE 802.11a standards.

The channel condition that determines what modulation scheme should be used at the SU's transceiver is divided into N^j_{SNR} states according to the SNR. Under each state, the modulation schemes are set to make sure that the average packet error rates are equal to a preset packet error rate P_{target}. Thus, the boundary of these states in terms of SNR can be written as:

$$\Gamma_n = \frac{1}{g_n} \ln\left(\frac{a_n}{P_{\text{target}}}\right), \tag{5.3}$$

where $\Gamma_n \in \{\Gamma_0, \Gamma_1, \ldots, \Gamma_{N^j_{\text{SNR}}}\}$ are the SNR boundaries of each state. $\Gamma_0 = 0$ and $\Gamma_{N^j_{\text{SNR}}+1} = \infty$ are set and the $N^j_{\text{SNR}} + 1$ SNR boundaries of the N^j_{SNR} states can be obtained. When the SNR at an SU's receiver is low ($\gamma \in (\Gamma_0, \Gamma_1)$), the SU cannot make any transmission and when the SNR satisfies $\gamma \in (\Gamma_n, \Gamma_{n+1}), n \in \{1, 2, \ldots, N^j_{\text{SNR}} - 1\}$, mode n is used in the transmission.

The model of the way in which the channel condition state evolves from time slot to time slot is assumed to be Rayleigh: the SNR is exponentially distributed with the probability density function:

$$p(\gamma^j_i) = \frac{1}{\bar{\gamma}^j_i} \exp\left(-\frac{\gamma^j_i}{\bar{\gamma}^j_i}\right), \quad \gamma^j_i \geq 0, \tag{5.4}$$

where γ^j_i is the SNR at the i-th SU on the j-th channel and $\bar{\gamma}^j_i$ is the average SNR. The steady state probability of i-th SU on the j-th channel in state k can thus be calculated as below:

$$\pi^j_{\text{CSi}}(k) = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma^j_i)\, \mathrm{d}x. \tag{5.5}$$

Let $\boldsymbol{P}^j_{\text{CSi}} = \begin{pmatrix} p_{0,0} & \cdots & p_{0,N^j_{\text{SNR}}-1} \\ \vdots & \ddots & \vdots \\ p_{N^j_{\text{SNR}}-1,0} & \cdots & p_{N^j_{\text{SNR}}-1,N^j_{\text{SNR}}-1} \end{pmatrix}$ denote the state transition probability matrix of the i-th SU on the j-th channel, where $p_{x,y}$ is the transition probability to the state y from the state x. Using the method in [67], the transition probabilities between the adjacent states can be calculated as below:

$$p_{k,k+1} \approx \frac{f_m T_p}{\pi^j_{\text{CSi}}(k)} \sqrt{\frac{2\pi\Gamma_{k+1}}{\bar{\gamma}^j_i}} \exp\left(-\frac{\Gamma_{k+1}}{\bar{\gamma}^j_i}\right), \tag{5.6}$$

Table 5.1. Transmission modes with convolutionally coded modulation

	Mode 1	Mode 2
Modulation	BPSK	QPSK
Coding rate R_c	1/2	1/2
Rate (bits/symbol)	0.5	1
a_n	274.7229	90.2514
g_n	7.9932	3.4998
Γ_n(dB)	-1.5331	1.0942

$$p_{k,k-1} \approx \frac{f_m T_p}{\pi^j_{\mathrm{CSi}}(k)} \sqrt{\frac{2\pi\Gamma_k}{\bar{\gamma}^j_i}} \exp\left(-\frac{\Gamma_k}{\bar{\gamma}^j_i}\right), \tag{5.7}$$

where f_m is the maximum Doppler frequency of the mobile station. The transition is assumed to occur only between adjacent states ($p_{j,k} = 0$, for any $|j-k| > 1$, and $p_{k,k} = 1 - p_{k,k+1} - p_{k,k-1}$). With Equations (5.3) to (5.7), the transition probability matrix $\boldsymbol{P}^j_{\mathrm{CSi}}$ can be obtained. Let $\boldsymbol{V}_{\mathrm{rate}} = \{V_{\mathrm{rate}}(k)|k \in \{0,\cdots,N^j_{\mathrm{SNR}}\}\}$ represent the number of packets that can be transmitted in one time slot when the channel condition state is k. The settings of the parameters above are listed in Table 5.1 and $\boldsymbol{V}_{\mathrm{rate}} = \{0,1,2\}$.

As the error recovery method, a stop-n-wait ARQ is applied. The average packet error rate is Per. After data transmission, the packets arriving during the current time slot will enter the SU buffer. If the total number of packets exceeds the buffer size, the overflow packets will be rejected.

5.3.3.1 Queueing model

1 ARRIVAL PROCESS The number of arrival packets at each SU during one time slot is assumed to follow a Batch Bernoulli process [46]. A probability vector $\boldsymbol{\alpha}_i$ describes the process of the i-th SU as follows:

$$\boldsymbol{\alpha}_i = (\alpha_i(0), \alpha_i(1), \cdots, \alpha_i(v_i)), \tag{5.8}$$

where $\alpha_i(j)$ is the probability that j packets arrives in one time slot at the i-th SU and v_i is the maximum number of packets that can arrive at the i-th SU during one time slot.

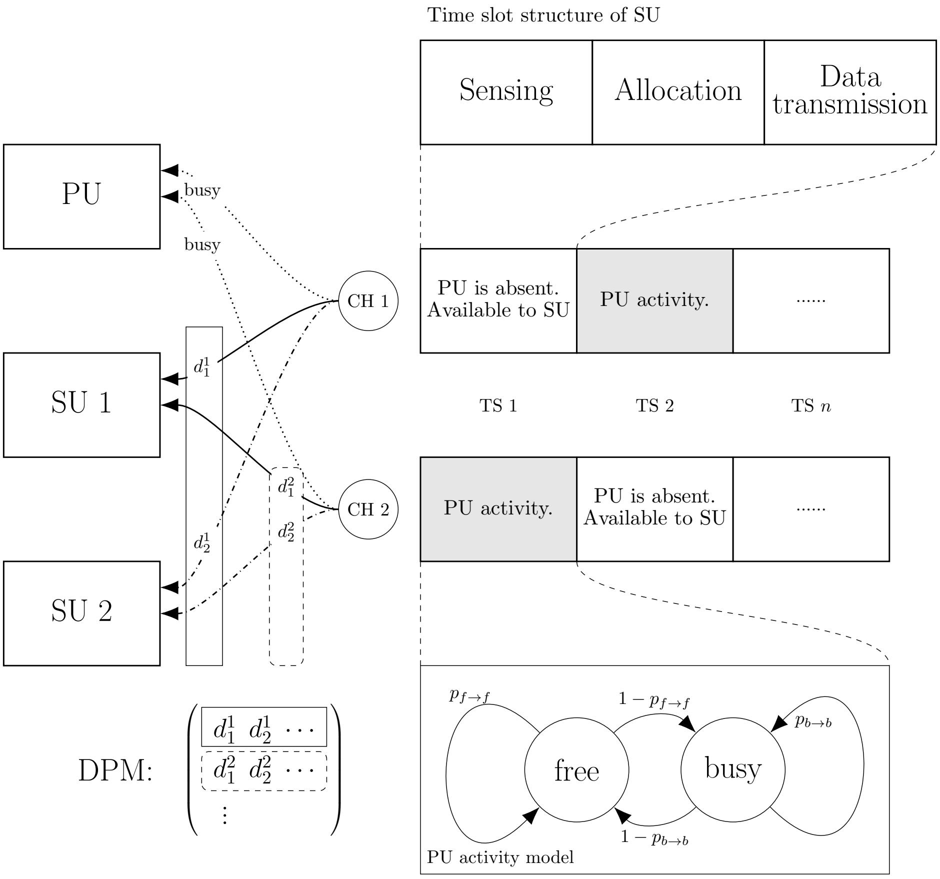


Figure 5.2. Illustration of DPM, time slot structure of SU system and the PU activity model.

2 QUEUEING MODEL With the analysis above, the queueing analytical model of the SU system can be built. In order to make the model capable of evaluating the performance of the SU system comprehensively, the setup of the state space of the system must include as many details as possible about the entire SU system, such as the states of the number of packets in the buffer of each SU, the states of PU activities on each channel and the state of channel conditions. The states of the number of packets in the buffer of each SU are denoted by:

$$\boldsymbol{b} \in \mathcal{B} = \{(b_1, b_2, \cdots, b_M) | b_i \in \{0, 1, \cdots, K_i\}, i \in [1, M]\}, \tag{5.9}$$

where $b_i \in \{0, 1, \cdots, K_i\}$ is the number of packets in the buffer of the i-th SU. The PU occupancy state of all J channels at the t-th time slot is represented by:

$$\boldsymbol{o} \in \mathcal{O} = \left\{ \left(O^1, O^2, \ldots, O^J\right) | O^j \in \{0, 1\} \right\}. \tag{5.10}$$

where O^j represents the PU activity state of the j-th channel at the t-th time slot. The channel condition state of J channels to M SUs is denoted by:

$$\boldsymbol{c} \in \mathcal{C} = \{(c_1, c_2, \cdots, c_M) \mid c_i \in [1, N_{\text{SNR}}(i)]\}, \tag{5.11}$$

where c_i is the channel condition state of the PU channels, and $N_{\text{SNR}}(i)$ is the number of channel condition states of the i-th SU as introduced in Section 5.3.3.

Finally, the state space of the system can be denoted by:

$$\Phi \triangleq \{(\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}) | \boldsymbol{b} \in \mathcal{B}, \boldsymbol{o} \in \mathcal{O}, \boldsymbol{c} \in \mathcal{C}\}. \tag{5.12}$$

Then, the transition matrix of the system $\boldsymbol{P}(\boldsymbol{D})$ given any DPM $\boldsymbol{D}$ can be obtained. The derivation of $\boldsymbol{P}(\boldsymbol{D})$ is given as below.

The state space of the SU system is denoted by:

$$\Phi \triangleq \{\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}\}. \tag{5.13}$$

The transition matrix between all the channel states $\{\boldsymbol{o}, \boldsymbol{c}\}$ is denoted by matrix $\boldsymbol{B}$ and $\boldsymbol{B}$ is obtained as:

$$\boldsymbol{B} = (\boldsymbol{P}_{\text{PU}}^1 \otimes \cdots \otimes \boldsymbol{P}_{\text{PU}}^J) \otimes \boldsymbol{P}_{\text{CS},1} \otimes \cdots \otimes \boldsymbol{P}_{\text{CS},M}, \tag{5.14}$$

where $\boldsymbol{P}_{\text{CS},i}$ is the channel condition state transition probability matrix of the i-th SU in Section 5.3.3.

First, the distribution of the number of packets that can be transmitted in one time slot should be derived. The distribution of transmitted packets of each SU through each channel, given the channel condition state $\boldsymbol{c}$ according to the AMC settings in Section 5.3.3, can be obtained. The number of transmission attempts that the i-th SU can make through j-th channel is denoted as C_i^j.

$$C_i^j = \begin{cases} V_{\text{rate}}(c_i) & \text{for } O^j \neq 0 \\ 0 & \text{for } O^j = 0 \end{cases}, \tag{5.15}$$

From the ARQ settings, the probability that x packets can be transmitted is calculated as follows:

$$\mu(x) = \begin{cases} \binom{C_i^j}{x}(1-\text{Per})^x & \text{for } 0 \leq x \leq C_i^j \\ \text{Per}^{C_i^j} & \text{for } x = 0 \end{cases}, \tag{5.16}$$

where $\binom{C_i^j}{x}$ is C_i^j choose x. Here, $\boldsymbol{A} = \{a_1, \cdots, a_J \mid a_j \in \{0, 1, \cdots, M\}\}$ is used to represent the channel allocation results. $a_j = 0$ represents the j-th channel occupied by the PU, and $a_j = i$ represents the j-th channel allocated to the i-th SU. Then, the distribution of number of packets that can be transmitted by the i-th SU through the j-th channel can be calculated as given below.

$$\boldsymbol{\mu}_i^j = \begin{cases} \{\mu_i^j(0), \ldots, \mu_i^j(C_i^j)\} & \text{for } a_j = i \\ \{0\} & \text{otherwise} \end{cases}. \tag{5.17}$$

The distribution of number of packets that can be transmitted by the i-th SU given the channel state $\{\boldsymbol{o}, \boldsymbol{c}\}$ can be obtained as follows:

$$\boldsymbol{\mu}_i(\boldsymbol{A}, \boldsymbol{o}, \boldsymbol{c}) = \boldsymbol{\mu}_i^1 * \boldsymbol{\mu}_i^2 * \cdots * \boldsymbol{\mu}_i^J, \tag{5.18}$$

where $*$ is convolution.

Then, the transition matrix that describes that the number of packets in the i-th SU's buffer changes from b_i to b_i' under $\boldsymbol{A}$ can be derived. First, a $(1+V) \times (1+C)$ auxiliary matrix that can select every possible arrival process and service process making every possible i-th SU's buffer change is constructed. Here, V is the maximum number of arrival packets and C is the maximum number of transmitted packets in one time slot. The $(1+v, 1+c)$-th element of the matrix is derived as follows:

$$\Gamma_{v,c}(i, i') = \begin{cases} 1 & \text{when } i' = \min(v + i - \min(i, c), K) \\ 0 & \text{otherwise} \end{cases}. \tag{5.19}$$

Thus, the transition matrix that describes every possible changes of the i-th SU's buffer can be written as follows:

$$\boldsymbol{P}(b_i, b_i', \boldsymbol{A}) = \begin{bmatrix} \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(0,0) \boldsymbol{\mu}_i(\boldsymbol{A}) \cdots \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(0,K) \boldsymbol{\mu}_i(\boldsymbol{A}) \\ \vdots \qquad \ddots \qquad \vdots \\ \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(K,0) \boldsymbol{\mu}_i(\boldsymbol{A}) \cdots \boldsymbol{\alpha}_i \boldsymbol{\Gamma}(K,K) \boldsymbol{\mu}_i(\boldsymbol{A}) \end{bmatrix}, \tag{5.20}$$

Set $\boldsymbol{b} = \{b_1, b_2, \cdots, b_M\}$ to denote the number of packets in each SU's buffer at the current time slot, and $\boldsymbol{b}' = \{b_1', b_2', \cdots, b_M'\}$ to denote the number of packets at the next time slot. Then, a coefficient matrix given $\boldsymbol{b}$, $\boldsymbol{b}'$ and $\boldsymbol{A}$ can be obtained:

$$\boldsymbol{T}(\boldsymbol{b},\boldsymbol{b}',\boldsymbol{A}) = \boldsymbol{P}(b_1,b_1',\boldsymbol{A}) \otimes \boldsymbol{P}(b_2,b_2',\boldsymbol{A}) \otimes \cdots \otimes \boldsymbol{P}(b_M,b_M',\boldsymbol{A}). \tag{5.21}$$

Given the DPM $\boldsymbol{D}$ and the PU occupancy state $\boldsymbol{o}$, the coefficient matrix can be obtained as below.

$$\boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\boldsymbol{D},\boldsymbol{o}) = \sum_{\mathrm{A}\in\mathcal{A}} \left(\prod_{j=1}^{J} d^j_{a_j} \boldsymbol{T}(\boldsymbol{b},\boldsymbol{b}',\boldsymbol{A}) \right), \tag{5.22}$$

where d_i^j is from the DPM $\boldsymbol{D}$ and $\mathcal{A}$ is the set of all possible allocation results. All the coefficients according to the index of channel condition states are listed, and are extended to a matrix with the same dimensions as $\boldsymbol{B}$:

$$\boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\boldsymbol{D},\boldsymbol{o}(1)) \\ \vdots \\ \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\boldsymbol{D},\boldsymbol{o}(\phi)) \\ \vdots \\ \boldsymbol{T}_{\boldsymbol{b},\boldsymbol{b}'}(\boldsymbol{D},\boldsymbol{o}(N_{\mathrm{CS}})) \end{bmatrix} \otimes \boldsymbol{1}_{1\times(\prod_{i=1}^{M}(K_i+1)\cdot N_{\mathrm{CS}})}, \tag{5.23}$$

where $\boldsymbol{o}(\phi)$ is the occupancy state that can be obtained given the system states ϕ, N_{CS} is the total number of system states. To enumerate all coefficient matrices, β_{index} is mapped to any $\{b_1,b_2,\cdots,b_M\}$ using the following equation:

$$\mathrm{index} = \sum_{i=1}^{M} b_i \times \mathrm{base}_{i-1}, \tag{5.24}$$

where $\mathrm{base}_i = \prod_{j=i}^{M}(K_i+1)$.

The transition matrix is divided into $\prod_{i=1}^{M}(K_i+1) \times \prod_{i=1}^{M}(K_i+1)$ blocks and obtained as follows:

$$\boldsymbol{P}(\boldsymbol{D}) = \begin{bmatrix} \boldsymbol{T}_{\beta_0,\beta_0} \circ \boldsymbol{B} \ldots \boldsymbol{T}_{\beta_0,\beta_B} \circ \boldsymbol{B} \\ \vdots \quad \ddots \quad \vdots \\ \boldsymbol{T}_{\beta_B,\beta_0} \circ \boldsymbol{B} \ldots \boldsymbol{T}_{\beta_B,\beta_B} \circ \boldsymbol{B} \end{bmatrix}, \tag{5.25}$$

where $\circ$ is the entrywise product.

5.3.4 Performance evaluation and optimization method

5.3.4.1 Performance measures derivation

The QR algorithm can be applied to the transition matrix $\boldsymbol{P}(\boldsymbol{D})$ to get the steady probability vector: $\boldsymbol{\pi}'$, which is a 1 by $\prod_{i=1}^{M}(K_i+1)\times 2^J\cdot\prod_{i=1}^{M}N_{\text{SNR}}(i)$ vector. Then, $\boldsymbol{\pi}'$ can be reorganized in the form of $\pi(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ by mapping a vector $\{\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}\}$ to any system state.

1 QUEUE LENGTH DISTRIBUTION Let $P(q_i=l)$ denote the probability that there are l packets waiting to be transmitted in the i-th SU's buffer, and it can be calculated as follows:

$$P(q_i=l)=\sum_{\boldsymbol{b}\in B_i(l)}\boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}), \tag{5.26}$$

where $B_i(l)=\{(b_1,\cdots,b_M)|b_i=l,b_j\in[0,K_j],j\neq i\}$, which represents the set of states of SU buffers when there are l packets in the i-th SU's buffer.

2 AVERAGE THROUGHPUT In this research, the expected number of packets that can be transmitted during one time slot is used to represent the average throughput of an SU. First, the distribution of number of packets the i-th SU can transmit during one time slot under each system state can be obtained, which is denoted by $\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$, as follows:

$$\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})=\sum_{A\in\boldsymbol{A}}d_i^jO^j\boldsymbol{\mu}_i(A,\boldsymbol{o},\boldsymbol{c}), \tag{5.27}$$

where $\boldsymbol{\mu}_i(A,\boldsymbol{o},\boldsymbol{c})$ is the distribution of number of packets that can be transmitted by the i-th SU when the allocation result is $A\in\boldsymbol{A}=\{a_1,\cdots,a_J\mid a_j\in\{0,1,\cdots,M\}\}$. $\boldsymbol{A}$ is the set of all possible allocation results: $a_j=0$ represents the j-th channel occupied by a PU, and $a_j=i$ represents the j-th channel allocated to the i-th SU. The PU occupancy state is $\boldsymbol{o}$ and the channel condition state is $\boldsymbol{c}$ from Equation (5.18). d_i^j is the element of the DPM $\boldsymbol{D}$ and O^j is the occupancy state of the j-th channel obtained from $\boldsymbol{o}$. Then, the average throughput under state $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ can be obtained:

$$\text{Thr}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})=\sum_{\beta}\min(\beta,b_i)\cdot t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},\beta), \tag{5.28}$$

where $t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},\beta)$ is the $(\beta+1)$-th element of $\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$, which is the probability that β transmissions have been made by the i-th SU under $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Thus, the average throughput of the i-th SU can be derived as:

$$\overline{\mathrm{Thr}}_i = \sum_{(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})\in\Phi} \mathrm{Thr}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) \cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}). \tag{5.29}$$

5.3.4.2 Weighted throughput

With the average throughput of every SU obtained by the framework, the overall analysis based on the average throughput can be now carried out easily. Here, the concept of weighted throughput is used: a parameter of weight is assigned to each SU that represents the level of importance. The weights of all SUs are represented by a set of positive numbers: $\boldsymbol{w} = \{(w_1, w_2, \cdots, w_M) \mid w_i \in \mathbb{R}^+\}$, where w_i is the weight of the throughput of the i-th SU. The total weighted throughput of the SU system given the DPM $\boldsymbol{D}$ can be calculated as follows:

$$\mathrm{W}(\boldsymbol{D}) = \sum_{i=1}^{M} w_i \cdot \overline{\mathrm{Thr}_i}. \tag{5.30}$$

5.3.4.3 Measure of delay

1 DISCUSSION ON CALCULATION OF DELAY The common approach to delay calculation is to make use of the arrival rates, service rates and the buffers' settings to reveal how long it takes a packet to be transmitted from its arrival. In the queueing analysis of the CR system, the delay is usually calculated as the time interval from when one tagged packet arrives to when it has been transmitted, which is actually the transmission time of the packets ahead of the tagged one. Thus, every details of transmission of the packets in front of the tagged packet needs to be considered step by step, which makes the analysis complicated, especially when there are multiple transmission rates, as when the AMC is used.

The calculation of delay above is accurate; however, it may be inappropriate to be applied to evaluate the CR system when computing efficiency is a concern. For instance, the measure of delay works as a reference for the CR system to allocate the PU channels at the beginning of a time slot. As a result, the longer it takes the CR system to calculate the measure of delay, the less time remains for the data transmission. To make the calculation of delay simpler, a different approach to the measure of delay is proposed: when a packet has been transmitted, there are a certain number of packets left in the buffer; thus, the time interval from when the packets arrive at the buffer until they leave can be used to represent the measure of delay of the transmitted packet. The idea behind this approach is that since the information on the service rates are already contained in the number of packets left in the buffer (which is the distribution of the queue length), it is possible to use the distribution of the queue length as replacement of the service rates to estimate the performance of delay.

2 CALCULATION OF THE MEASURE OF DELAY The measure of delay in this research is defined as the waiting time of a packet from the time slot when it arrives to the time slot when it leaves the buffer. At the time slot a packet is being transmitted, it can be assumed that there are x packets in the buffer. Thus, the estimation of the time interval during which x packets has arrived can be used to estimate the waiting time of the packet being transmitted. The actual waiting time t' consists of two parts. The first part is the time t during which x packets arrives and the second part is the time slots when continuous packet rejection happens when the buffer is full. It is obvious that when the buffer size of the SU in the model approaches to infinity, the rejection probability of packets converges to zero. Thus to trade accuracy for efficiency, the first part can be used solely to estimate the measure of delay. The error caused by the omission of the second part can be eliminated by increasing the buffer size in the model.

According to the discussion above, the probability that the packet arrival time t is greater than a preset length of time is used to estimate the measure of delay in this framework. This probability is named OP and the preset length of time is named delay reference. The OP is obtained as follows: First, the delay reference is represented by an integer d_{ref}, which is the number of time slots. For the i-th SU, when there are x packets in the buffer, the probability that arrival time t equals j time slots is:

$$P_i(t=j) = \sum_{y=1}^{v} \alpha_i(y) \cdot \text{Conv}(\boldsymbol{\alpha}_i, j-1, x-y), \tag{5.31}$$

where $\text{Conv}(\boldsymbol{\alpha}, i, j)$ is the j-th element of vector $\underbrace{\boldsymbol{\alpha} * \boldsymbol{\alpha} * \cdots * \boldsymbol{\alpha}}_{i\text{`}\boldsymbol{\alpha}\text{'}s}$, $*$ is the discrete convolution of vectors and $\boldsymbol{\alpha}_i$ is the vector that represents the arrival process of the i-th SU in Equation (5.8).

Finally, the OP, given the delay reference d_{ref} time slots, can be obtained:

$$\begin{aligned} P_i(d_{\text{ref}}) = {} & P(q_i=0)\alpha(0)^{d_{\text{ref}}} + \\ & \sum_{x=1}^{K_i} P(q_i=x) \cdot (1 - \sum_{i=1}^{d_{\text{ref}}} \sum_{y=1}^{v} \alpha_i(y) \cdot \text{Conv}(\boldsymbol{\alpha}_i, i-1, x-y)), \end{aligned} \tag{5.32}$$

where $P(q_i = l)$ is the probability that there are l packets in the i-th SU's buffer as in Equation (5.26).

5.4 OPTIMIZATION PROBLEM FORMULATION AND SOLUTION

There are two important concerns in the study of the RA problem in CRN. The first one is how to make the best use of the limited radio resource to satisfy the transmission demand of the SUs. The second one is how to guarantee the minimum requirement of each SU with the limited resource. For the first concern, to describe the extent of use of resource, it is crucial to have a flexible and configurable criterion that is adaptable to the variety of SU types. The concept of weighted throughput is one of the solutions. For the second concern, the delay is one of the requirements that is usually considered in the related research.

Using the proposed framework, the weighted throughput of the SU system and delay measures of each SU given any $\boldsymbol{D}$ can be obtained efficiently. Thereafter, with the two concerns discussed above, the optimal channel allocation under the criteria, including requirements of the weighted throughput and the delay, can be investigated.

In this part, the author has set up an optimization objective with the requirements of the weighted throughput and the delay as follows: Find a DPM $\boldsymbol{D}$ that can maximize the overall weighted throughput of the whole SU system, while the measure of delay of every SU, which is represented by OP in Equation (5.32), is no greater than a threshold P_{delay}. The optimization problem can be written in the following form:

$$\begin{aligned} \underset{\boldsymbol{D}}{\text{maximize}} \quad & W(\boldsymbol{D}) \\ \text{subject to} \quad & P_i(d_{\text{ref}}) < P_{\text{delay}}, \text{for } \forall i \in [1, M]. \end{aligned} \tag{5.33}$$

With the same concern of computing efficiency as mentioned in the delay calculation part, the SU system needs an efficient method to obtain an approximate solution of DPM under certain precisions rather than a closed form solution of the DPM. To find an approximate DPM solution to this optimization problem, a modified hill climbing algorithm [123, 124] is applied. The detail of the algorithm is

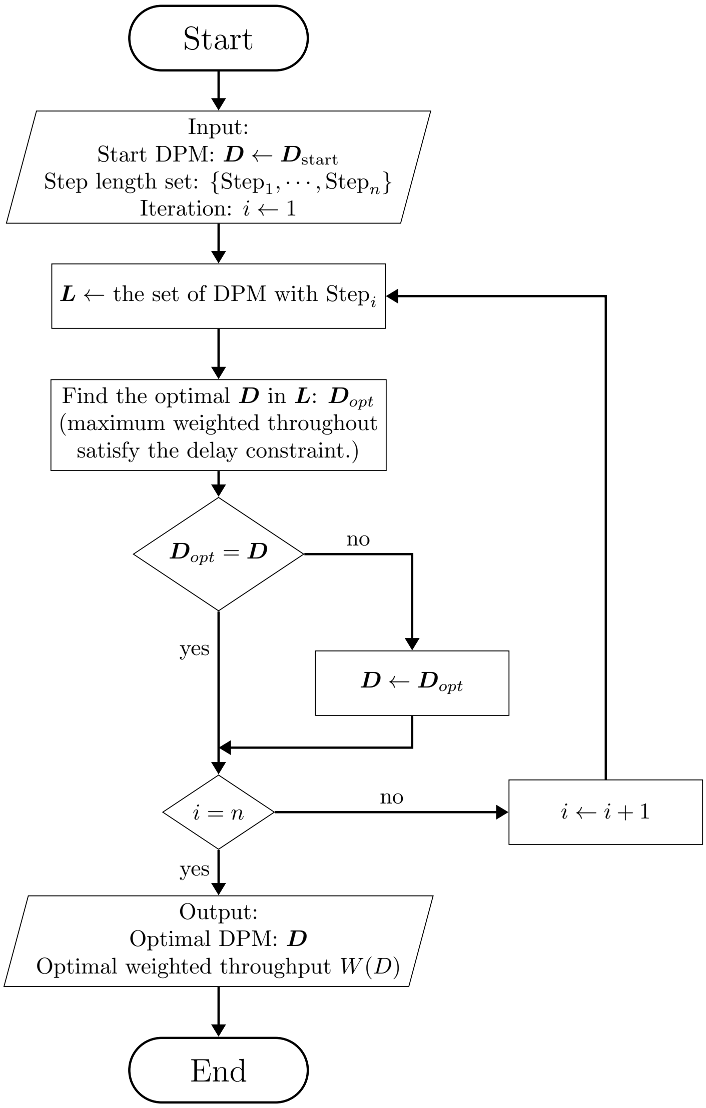


Figure 5.3. The flow chart of the hill climbing algorithm.

shown in Algorithm 5. The process of the algorithm is shown in Figure 5.3. First, set a matrix of DPM as the start DPM and set a step as the first element in a preset step length set $\{\text{Step}_1, \cdots, \text{Step}_n\}$. Then, search every element of the DPM using the step to find the best weighted throughput that satisfies the delay constraint meanwhile. Then, reset the step value as the next element in the step length set, and repeat the search until all the values in the step length set have been searched.

Algorithm 5 Discrete space hill climbing algorithm (Part I)

Input:
Weighted throughput function: W
Delay evaluation function of SU i: P_i
Delay reference: d_{ref}
Set of step lengths: $\{\text{Step}_1, \cdots, \text{Step}_n\}$
Initial DPM: $\boldsymbol{D}_{\text{start}} = \begin{pmatrix} d_1^1 \cdots d_M^1 \\ \vdots \ \ddots \ \vdots \\ d_1^J \cdots d_M^J \end{pmatrix}$

Output: optimal weighted throughput: W_{opt}
optimal DPM: $\boldsymbol{D}_{\text{opt}}$

$\boldsymbol{D} \leftarrow \boldsymbol{D}_{\text{start}}$;
$i \leftarrow 1$
Step $\leftarrow$ Step$_i$;
repeat
  $W_{\text{current}} \leftarrow W(\boldsymbol{D})$;
  Let L be a set of DPM;
  for j in $[1, J-1]$ **do**
    for m in $[1, M]$ **do**
      if $d_m^j + \text{Step} \in [0,1]$ **then**
        $d_m^j \leftarrow d_m^j + \text{Step}$
        add $\boldsymbol{D}$ to L
      end if
      if $d_m^j - \text{Step} \in [0,1]$ **then**
        $d_m^j \leftarrow d_m^j - \text{Step}$
        add $\boldsymbol{D}$ to L
      end if
    end for
  end for
  findOpt $\leftarrow$ False;
  for all $\boldsymbol{D}$ in L **do**
    $T(\boldsymbol{D}) \leftarrow$ maximun in P_i of all SU given $\boldsymbol{D}$
    if $W(\boldsymbol{D}) > W_{\text{current}}$ AND $T(\boldsymbol{D}) < d_{\text{ref}}$ **then**
      $\boldsymbol{D}_{\text{next}} \leftarrow \boldsymbol{D}$;
      $W_{\text{current}} \leftarrow W(\boldsymbol{D})$;
      findOpt $\leftarrow$ True;
    end if
  end for

Algorithm 5 Discrete space hill climbing algorithm (Part II)

```
    if findOpt then
        D ← D_next;
    end if
    if Step = Step_n then
        endCondition ← True;
    else
        i ← i + 1;
        Step ← Step_i;
    end if
until endCondition = True;
W_opt ← W_current;
D_opt ← D;
```

5.5 NUMERICAL RESULTS AND ANALYSIS

In the following section, the framework that has been introduced is implemented to evaluate the performance of the SU system and acquire qualified DPM from Equation (5.33) under various system configurations. To make the results simple and without loss of generality, a model with two users and two PU channels is implemented. The default settings of the parameters are as follows:

- Channel condition: $P_{\text{target}} = 0.01$, $f_m = 10\,\text{Hz}$
- PU activity model: $p^1_{\text{b}\to\text{b}} = p^2_{\text{b}\to\text{b}} = p^1_{\text{f}\to\text{f}} = p^2_{\text{f}\to\text{f}} = 0.5$
- Arrival process: $\alpha_1 = \alpha_2 = \{0.5, 0.5\}$
- SU buffer size: $K_1 = 6$, $K_2 = 6$
- ARQ setting: $\text{Per} = 0.01$.

5.5.1 Throughput and weighted throughput

The average throughput of both SUs can be obtained simultaneously and the weighted throughput of the SU system under any weights can also be obtained directly with only one run of calculation using this framework. The framework makes it convenient to compare the throughputs between different SUs and analyze the overall performance of the SUs. The comparison of the throughputs of both SUs and the analysis of the weighted throughput with different weight settings under different channel conditions are shown in Figure 5.4. In Figure 5.4(a)(b), the influence of DPM on the throughput of

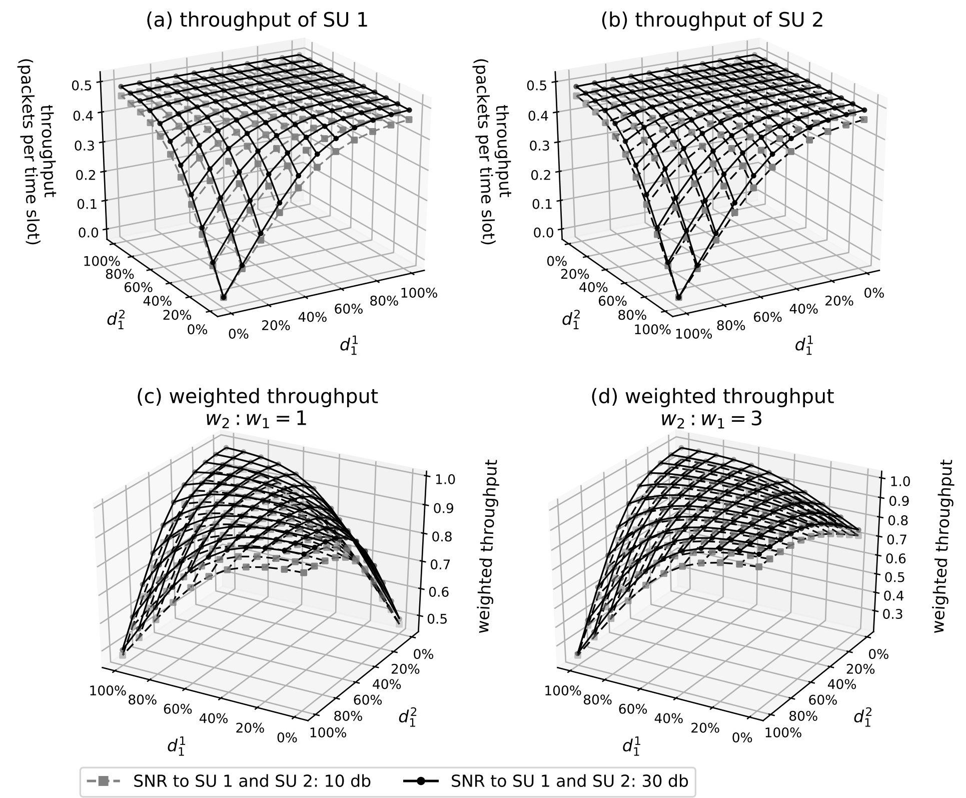


Figure 5.4. The throughput of each SU and weighted throughput function of the SU system versus the DPM. d_i^j is the probability that the j-th channel is allocated to the i-th SU.

both SUs is shown. It is evident that the improvement of throughput caused by increased distribution probability is greater when the DPM is lower. However, when the DPM to an SU is high (greater than 50%), the further increase of the DPM dose not improve the throughput significantly. The effect of this trend on the weighted throughput is shown in Figure 5.4(c) and (d). Figure 5.4(c) shows the weighted throughput function versus the DPM when the throughputs of both SUs have the same weights ($w_1 = w_2$), while Figure 5.4(d) reflects the weights of SU 2 being greater than the SU 1 ($w_2 : w_1 = 3$). All the trends are identical when the SNR conditions (in terms of SNR) of SUs vary.

5.5.2 Probability of overtime

In this section, the OP that is the estimation of the measure of delay in this research is analyzed. Figure 5.5 illustrates that the value of the OP versus DPM under different delay references. The OP of SUs with different SNRs versus the DPM is shown in Figure 5.6; it is obvious that the SU with a higher SNR has a lower OP. The advantages of the proposed framework are as follows. First, the OP is obtained only from the queue length distribution and the references and the queue length distribution is

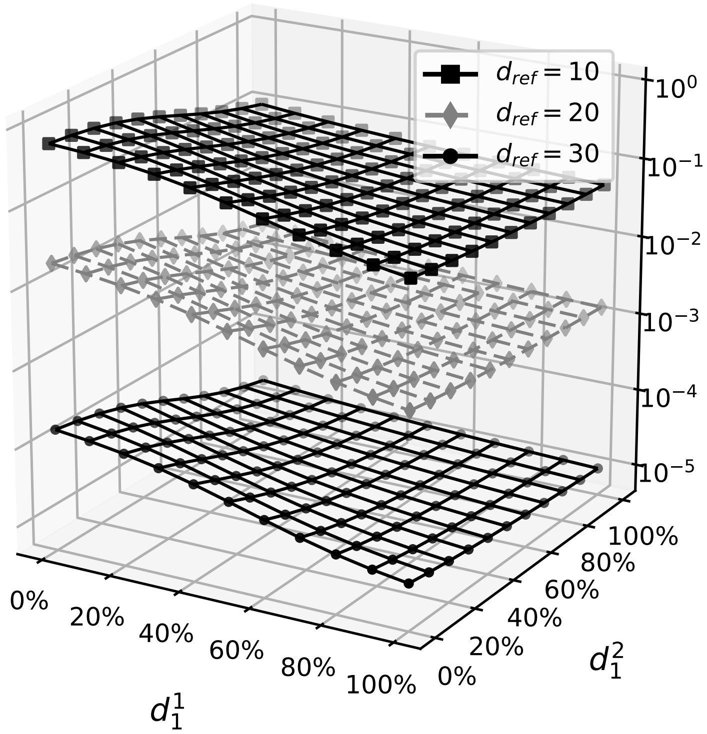
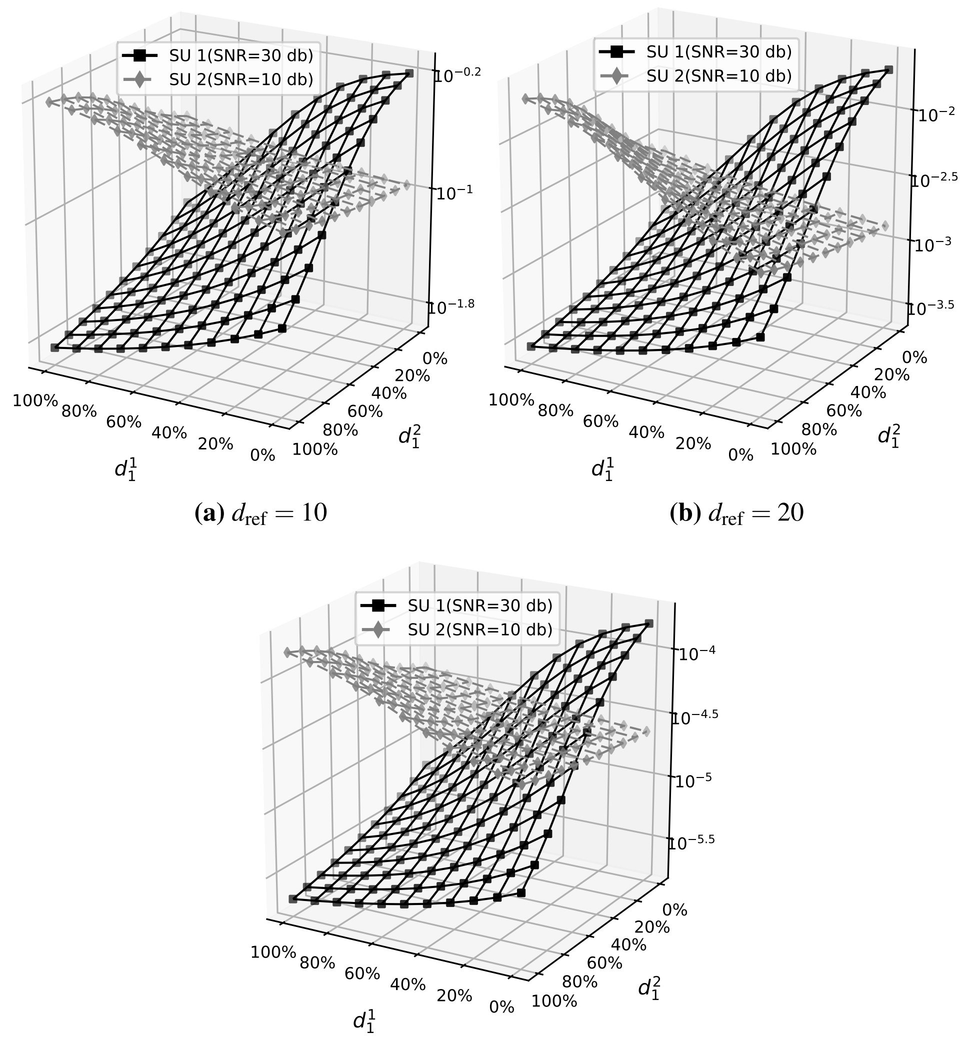


Figure 5.5. The probability of overtime versus DPM under different d_{ref}.

independent of the reference. As long as the queue length distribution under certain system settings and environmental parameters can be obtained, the OP with any delay reference can be obtained with simple calculations. Second, since the calculation of the measure of delay is an efficient approximation, the computing time is significantly reduced compared to other delay calculation methods. A comparison of run time statistics between the method employed in this study (Method I) and an example of other methods in the related literatures (Method II) is shown in Figure 5.7. The details of the Method II are given as below.

Method II to calculate the measure of delay uses a similar idea as that used in former research, such as [120]. The measure of delay is represented by the expected time to transmit all packets that are in the buffer when the packet that is considered arrives. Since the AMC is applied and the transmission rate varies, the state space is extended with another state f that represents the number of packets that have already been transmitted by a selected SU, to trace the transmission process of the packets ahead of the tagged one. The extended system state is $\Phi' \triangleq \{\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}, f\}$. The delay of a tagged packet can be interpreted as the first passage time from the system state when this packet arrives to the system state when all packets ahead are transmitted (number of packets in the i-th SU buffer $b_i = 0$ or f equals the number of packets in the buffer).



(a) $d_{\text{ref}} = 10$

(b) $d_{\text{ref}} = 20$

(c) $d_{\text{ref}} = 30$

Figure 5.6. The probability of overtime of SUs with different SNRs versus DPM under different $d_{\text{ref}} \in \{10, 20, 30\}$. The SNR of SU 1 is 30 dB, the SNR of SU 2 is 10 dB.

To calculate the first passage time above, the transition matrix is modified based on $\boldsymbol{P}(\boldsymbol{D})$ (to make it concise the $\boldsymbol{D}$ is omitted in the equations) given that there are F packets in the selected SU's buffer. First, an extending band matrix is built given $\{\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}\}$ and $\boldsymbol{t}_i$ obtained as in Equation (5.27):

$$\boldsymbol{P}_{\text{ex}}(\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}, f) = \begin{bmatrix} P_{0,0} \cdots P_{0,f} \\ \vdots \quad \ddots \quad \vdots \\ P_{f,0} \cdots P_{f,f} \end{bmatrix}, \tag{5.34}$$

where

$$P_{j,j'} = \begin{cases} t_i(\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}, j'-j) & \text{for } j'-j \in [0, \max(\boldsymbol{t}_i(\boldsymbol{b}, \boldsymbol{o}, \boldsymbol{c}))] \\ 0 & \text{otherwise} \end{cases},$$

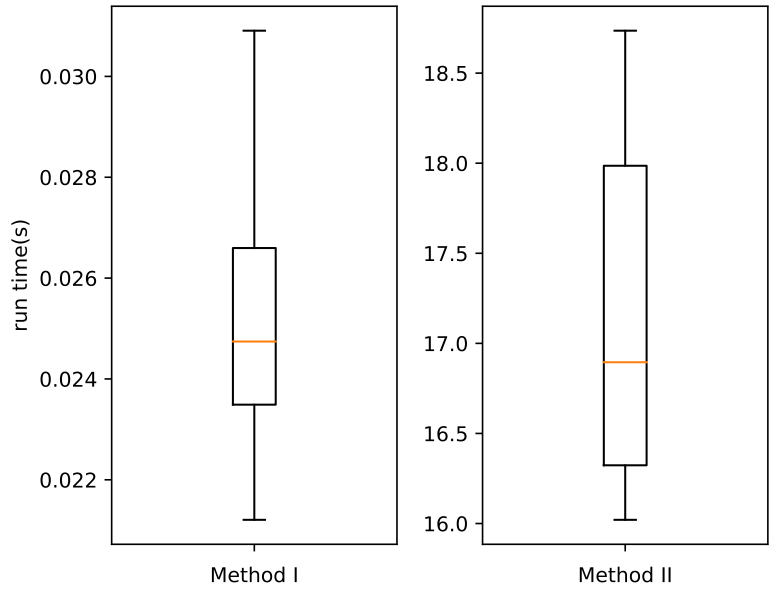


Figure 5.7. Comparison of delay computing time statistics.

where $\max(\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}))$ is the maximum number of packets that can be transmitted when the system state is $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Then, the $\boldsymbol{P}$ is extended row by row as follows:

$$\boldsymbol{P}'_{\text{delay}}(f) = \begin{bmatrix} \boldsymbol{P}(\phi(0)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(0), f) \\ \vdots \\ \boldsymbol{P}(\phi(R_m)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(R_m), f) \end{bmatrix}, \tag{5.35}$$

where $R_m = \prod_{i=1}^{M}(K_i+1)\cdot N_{\text{CS}}$ is the size of $\boldsymbol{P}$, and $P(\phi(r))$ is the r-th row of matrix $\boldsymbol{P}$. $\phi(i)$ represents the system state of the i-th row in $\boldsymbol{P}$. The absorbing states are when all the packets in the buffer are transmitted ($b_i = 0$ or $f = F$), then they are removed from $\boldsymbol{P}'_{\text{delay}}$ and $\boldsymbol{P}_{\text{delay}}$ is obtained. The average first passage time to the absorbing states under each state is denoted as $\boldsymbol{S}$, which is a $(K_s \cdot \prod_{i\neq s}(K_i+1)\cdot N_{\text{CS}})$-by-1 column vector ($s$ is the label of the selected SU). Then, $\boldsymbol{S}$ can be solved from the following equation system:

$$(\boldsymbol{I} - \boldsymbol{P}_{\text{delay}}) \cdot \boldsymbol{S} = \boldsymbol{1}, \tag{5.36}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{1}$ is a column vector whose elements are all 1 and its size is the same as $\boldsymbol{S}$. Let $\boldsymbol{S}(F)$ be the solution of $\boldsymbol{S}$ given F. Thus the average first passage time from the state of F

packets in the buffer to the empty buffer state is:

$$d(F) = \sum_{\boldsymbol{b} \in B_s(F)} S(F,\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}) \cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}), \tag{5.37}$$

where $B_s(F)$ is the set of all states of SU buffers when the selected SU's buffer has F packets in it and $S(F,\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ is the elements in $\boldsymbol{S}$ when the system state is $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Finally the average delay of the i-th SU can be obtained as:

$$\overline{\text{Dly}} = \sum_{t=1}^{K_i} d(t). \tag{5.38}$$

In this section, the method to calculate the measure of delay using a similar idea as that used in earlier research such as [120] is introduced. The measure of delay is represented by the expected time to transmit all packets that are in the buffer when the packet that is considered arrives. Since the AMC is applied and the transmission rate varies, the state space is extended with another state f that represents the number of packets that have already been transmitted by a selected SU to trace the transmission process of the packets ahead of the tagged one. The extended system state is $\Phi' \triangleq \{\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},f\}$. The delay of a tagged packet can be interpreted as the first passage time from the system state when this packet arrives at the system state that all packets ahead are transmitted (number of packets in the i-th SU buffer $b_i = 0$ or f equals the number of packets in the buffer).

To calculate the first passage time above, the transition matrix is modified based on $\boldsymbol{P}(\boldsymbol{D})$ (to make it concise the $\boldsymbol{D}$ is omitted in the equations) given that there are F packets in the selected SU's buffer. First, an extending band matrix is built given $\{\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}\}$ and $\boldsymbol{t}_i$ obtained as in Equation (5.27):

$$\boldsymbol{P}_{\text{ex}}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},f) = \begin{bmatrix} P_{0,0} \cdots P_{0,f} \\ \vdots \ \ddots \ \vdots \\ P_{f,0} \cdots P_{f,f} \end{bmatrix}, \tag{5.39}$$

where

$$P_{j,j'} = \begin{cases} t_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c},j'-j) & \text{for } j'-j \in [0, \max(\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}))] \\ 0 & \text{otherwise} \end{cases},$$

where $\max(\boldsymbol{t}_i(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}))$ is the maximum number of packets that can be transmitted when the system state is $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Then the $\boldsymbol{P}$ is extended row by row as follows:

$$\boldsymbol{P}'_{\text{delay}}(f) = \begin{bmatrix} \boldsymbol{P}(\phi(0)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(0),f) \\ \vdots \\ \boldsymbol{P}(\phi(R_m)) \otimes \boldsymbol{P}_{\text{ex}}(\phi(R_m),f) \end{bmatrix}, \tag{5.40}$$

where $R_m = \prod_{i=1}^{M}(K_i+1)\cdot N_{\text{CS}}$ is the size of $\boldsymbol{P}$, and $P(\phi(r))$ is the r-th row of matrix $\boldsymbol{P}$. $\phi(i)$ represents the system state of the i-th row in $\boldsymbol{P}$. The absorbing states are when all the packets in the buffer are transmitted ($b_i = 0$ or $f = F$), then they are removed from $\boldsymbol{P}'_{\text{delay}}$ and $\boldsymbol{P}_{\text{delay}}$ is obtained. The average first passage time to the absorbing states under each state is denoted as $\boldsymbol{S}$, which is a $(K_s \cdot \prod_{i\neq s}(K_i+1)\cdot N_{\text{CS}})$-by-1 column vector ($s$ is the label of the selected SU). Then $\boldsymbol{S}$ can be solved from the following equation system:

$$(\boldsymbol{I} - \boldsymbol{P}_{\text{delay}})\cdot \boldsymbol{S} = \mathbf{1}, \tag{5.41}$$

where $\boldsymbol{I}$ is the identity matrix, $\mathbf{1}$ is a column vector whose elements are all 1 and size is the same as $\boldsymbol{S}$. Let $\boldsymbol{S}(F)$ be the solution of $\boldsymbol{S}$ given F. Thus the average first passage time from the state of F packets in the buffer to the empty buffer state is:

$$d(F) = \sum_{\boldsymbol{b}\in B_s(F)} S(F,\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})\cdot \boldsymbol{\pi}(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c}), \tag{5.42}$$

where $B_s(F)$ is the set of all states of SU buffers when the selected SU's buffer has F packets in it and $S(F,\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$ are the elements in $\boldsymbol{S}$ when the system state is $(\boldsymbol{b},\boldsymbol{o},\boldsymbol{c})$. Finally the average delay of the i-th SU can be obtained as:

$$\overline{\text{Dly}} = \sum_{t=1}^{K_i} d(t). \tag{5.43}$$

5.5.3 Optimization and discussion

After the weighted throughput and the OP have been obtained with this framework, the proposed hill climbing algorithm can be applied to find the optimal DPM that can achieve the best weighted throughput under delay constraints. An example of the optimization problem is illustrated in Figure 5.8. The objective of the algorithm is to find the best weighted throughput when the OPs of both SUs are feasible. The feasible area in Figure 5.8 is the area between the contour lines of delay constraint P_{delay}.

The processes of the algorithm under different settings are shown in Table 5.2 and Table 5.3. It is evident that if there are no delay constraints, the algorithm can converge to the solution rapidly and if there are delay constraints the rate of convergence becomes worse.

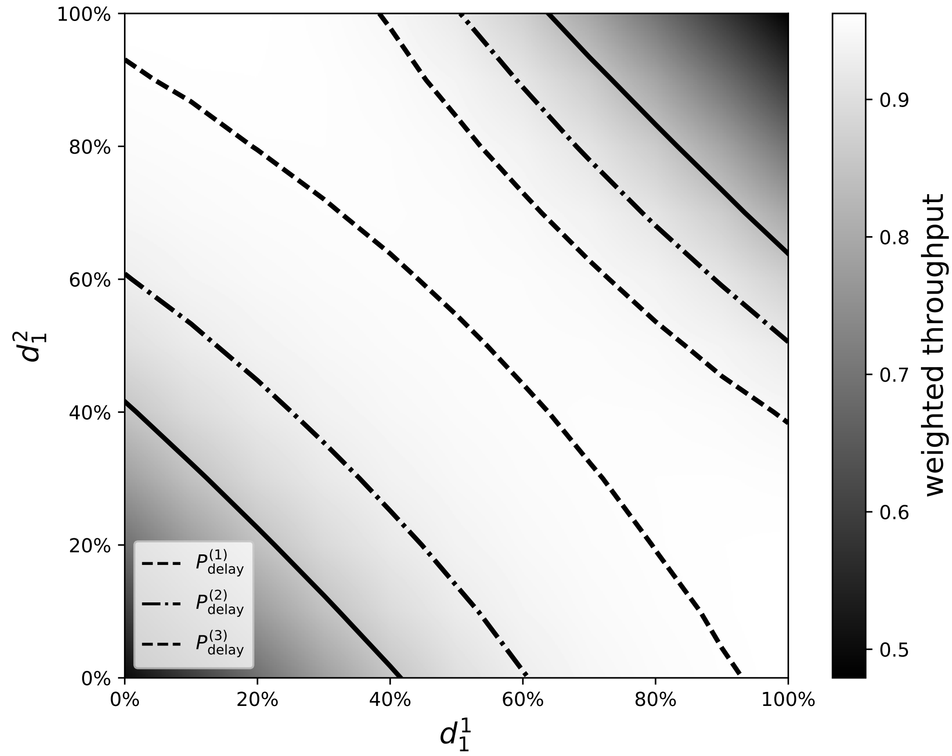


Figure 5.8. Illustration of the optimization problem. The settings are as follows: SNR of SU1 and SU2: 10 dB, $w_2 : w_1 = 10$, d_{ref}=20, $P^{(1)}_{\text{delay}} = 2\times10^{-3}$, $P^{(2)}_{\text{delay}} = 4\times10^{-3}$, $P^{(3)}_{\text{delay}} = 8\times10^{-3}$.

Table 5.2. Optimization process with no delay constraints

iterations	d_1^1	d_1^2	$W(\boldsymbol{D})$	OP 1	OP 2
start	0.5	0.5	1.127 183	5.136×10^{-3}	2.85×10^{-8}
1	0.4	0.4	1.138 202	6.441×10^{-3}	1.61×10^{-8}
2	0.3	0.5	1.139 135	6.421×10^{-3}	1.60×10^{-8}
3	0.2	0.6	1.141 875	6.362×10^{-3}	1.55×10^{-8}
4	0.1	0.7	1.146 248	6.265×10^{-3}	1.46×10^{-8}
5	0	0.8	1.151 963	6.133×10^{-3}	1.36×10^{-8}
6	0	0.9	1.156 396	5.398×10^{-3}	1.70×10^{-8}
7	0	0.91	1.156 426	5.337×10^{-3}	1.75×10^{-8}
8	0	0.909	1.156 427	5.343×10^{-3}	1.74×10^{-8}
Parameter Settings					
SNR of SU 1	5 dB	$\boldsymbol{\alpha}_1$	$\{0.5, 0.5\}$		
SNR of SU 2	30 dB	$\boldsymbol{\alpha}_2$	$\{0.2, 0.8\}$		
d_{ref}	20	P_{delay}	1		

The optimal DPM under different weights and delay constraints is shown in Figure 5.9. From Figure 5.9(a), it is clear that when there is no delay constraint ($P_{\text{delay}} = 1$), the channels need to be allocated more frequently to SU 2 to achieve better overall weighted throughput as the weight of SU

Table 5.3. Optimization process with delay constraints

iterations	d_1^1	d_1^2	$W(\boldsymbol{D})$	OP 1	OP 2
start	0.5	0.5	1.127183	5.136×10^{-3}	2.85×10^{-8}
1	0.4	0.6	1.128324	5.124×10^{-3}	2.83×10^{-8}
2	0.3	0.7	1.131703	5.088×10^{-3}	2.75×10^{-8}
3	0.2	0.8	1.137177	5.031×10^{-3}	2.62×10^{-8}
4	0.1	0.9	1.144493	4.955×10^{-3}	2.43×10^{-8}
5	0	1	1.153264	4.864×10^{-3}	2.20×10^{-8}
6	0	0.99	1.153931	4.910×10^{-3}	2.14×10^{-8}
7	0	0.98	1.154517	4.958×10^{-3}	2.09×10^{-8}
8	0	0.97	1.155022	5.007×10^{-3}	2.03×10^{-8}
9	0	0.96	1.155448	5.058×10^{-3}	1.98×10^{-8}
10	0	0.95	1.155796	5.111×10^{-3}	1.93×10^{-8}
11	0	0.94	1.156067	5.165×10^{-3}	1.88×10^{-8}
12	0	0.939	1.156090	5.170×10^{-3}	1.88×10^{-8}
13	0	0.938	1.156112	5.176×10^{-3}	1.87×10^{-8}
14	0	0.937	1.156134	5.181×10^{-3}	1.87×10^{-8}
15	0	0.936	1.156154	5.187×10^{-3}	1.86×10^{-8}
16	0	0.935	1.156174	5.192×10^{-3}	1.86×10^{-8}
17	0	0.934	1.156193	5.198×10^{-3}	1.85×10^{-8}
Parameter Settings					
SNR of SU 1	5 dB	$\boldsymbol{\alpha}_1$	$\{0.5, 0.5\}$		
SNR of SU 2	30 dB	$\boldsymbol{\alpha}_2$	$\{0.2, 0.8\}$		
d_{ref}	20	P_{delay}	5.2×10^{-3}		

2 increases. However, the decrease of the DPM of one SU (d_1^1 and d_1^2) makes its OP increase. Thus, when there are delay constraints, the increase in optimal DPM caused by the increase of weight reaches a critical point where the OP is equal to the delay constraint P_{delay}. The weight setting under which the delay constraints met is shown in Figure 5.9(b). The changes in optimal DPM with delay constraints are also shown in Figure 5.9(a).

5.6 CONCLUSION

In this chapter, a practical optimization objective of a multi-user multi-channel CRN is constructed as follows: Given all the environmental parameters and system settings, a way is examined to configure the channel allocation protocols to maximize the overall weighted throughput of the whole SU system while each SU is under certain delay constraints. This optimization problem is formulated and solved using the DPM mechanism and the queueing analytical framework based on DPM. Efforts are made to improve the accuracy and efficiency of the solution process. First, the proposed queueing analytical

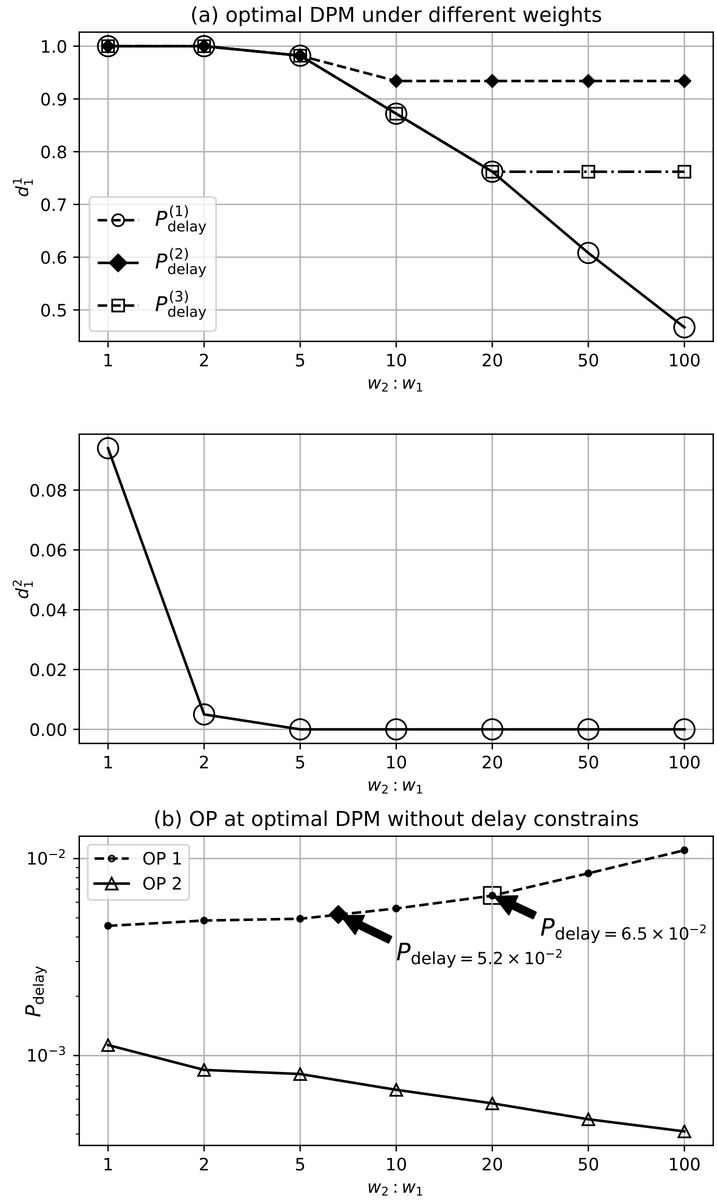


Figure 5.9. Optimal DPM versus weights under different delay constraints. $P_{\text{delay}}^{(1)} = 1$, $P_{\text{delay}}^{(2)} = 6.5 \times 10^{-3}$, $P_{\text{delay}}^{(3)} = 5.2 \times 10^{-3}$.

framework based on the DPM is able to obtain various performance metrics of each SU independently and simultaneously. Second, an approximation calculation of the measure of delay named OP is introduced to describe the delay constraints, which consume less computational resources than other measure of delay calculation methods. Third, a modified hill climbing algorithm is designed to obtained the numerical solution of the optimization problem under configurable precision.

The process of optimization problem formulation and solution above proves that the DPM mechanism and the queueing analytical framework based on the DPM are able to help manage RA in a multi-user multi-channel CRN to achieve comprehensive overall objectives under complicated constraints. Accordingly, the potential QoS that a CR system is able to provide under certain conditions can be revealed. Moreover, the performance evaluations and optimization solution enabled by the DPM mechanism can serve as a comprehensive and accurate reference to describe a single cell or cluster inside a CRN and help to analyze the inter-cell or inter-cluster behavior of the entire CRN.

CHAPTER 6 CONCLUSION

6.1 SUMMARY

CR technology, with its great potential to help resolve the conflict between the growing demand for wireless communication and the underutilization of the scarce radio spectrum resources, has attracted enormous attention in the wireless communication field. To achieve the promise of CR technology to solve the spectrum scarcity problem, RA in CRN is crucial. However, both the diversity of the users that a CR system probably serves and the variety of radio resources that a CR system is allowed to utilize make it a challenge to find a versatile mechanism that can guide the design of RA protocols and carry out evaluations of the protocols. The author has identified a knowledge gap where not enough work has been done on the design of the type of mechanisms mentioned above and consequently undertook the research study presented in this thesis. The findings and investigations presented in this thesis provide a practical and efficient solution to address the RA protocol design and evaluation problems in multi-user multi-channel CRN and furthermore, make it possible to explore the potential of CR technology.

The following concluding section summarizes the core ideas, contributions to knowledge and valuable findings according to the chapter structure of the thesis. After the presentation of the main contents of the thesis following the basic principles of technical reporting in the field of electronic and telecommunications engineering, recommendations are made for possible improvement in the future and extensive realistic considerations.

In Chapter one, a brief introduction to the research work was provided. The technical backgrounds and premises were stated and the principal objectives, based on the aim of the whole research work, were formulated. The definition of the problem that this thesis was devoted to solving was established, which was to design a mechanism to create and evaluate RA protocol in CRN to solve the challenge of making full use of limited radio resources, as well as a analytical framework to carry out the evaluations efficiently and accurately. The contributions realized by the work carried out during the research are listed, together with a list of publications based on the research.

In Chapter two, a comprehensive literature study on RA protocol in CRN was conducted. The common methodologies to model CRN and evaluate the performance metrics were investigated, followed by a summary of existing RA protocols based on the methodologies. Then, the chapter revealed the weakness in the principle of the protocols and suggested possible directions that can improve the performance of the CRN. The chapter is concluded by the research done to bridge the gaps in RA protocol in CRN.

In Chapter three, the essence of the entire research was proposed: A mechanism for channel allocation protocol in CRN, namely DPM. The two basic components of the DPM mechanism, the allocation indicator matrix and the effective duration, were introduced, followed by the explanation of the typical usages of the DPM. In the remaining part of the chapter, the DPM was implemented as a channel allocation protocol in a multichannel CRN and the queueing analysis was carried out accordingly to acquire the performance metrics using DPM and reference protocols. The flexibility and adaptability of the DPM as channel allocation protocol, together with the comparisons with existing protocols were demonstrated by the numerical results. The assertion that the DPM has advantages as a channel allocation protocol and that DPM has potential to provide references to determine the SU system settings to achieve the performance requirements are valid finally concludes this chapter.

In Chapter four, besides working as a channel allocation protocol, the usage of the DPM mechanism to help design and evaluate channel allocation protocols with specific objectives was demonstrated by an implementation of a protocol, namely MT. A powerful queueing analytical framework that is capable of obtaining the performance metrics of each CR user independently was established, based on DPM, to carry out the evaluations in a multi-user multi-channel CRN scenario. Through the framework, the MT protocol was analyzed and compared with reference protocols also aimed at maximizing the throughput of the CR system. Numerical results demonstrated the improvements on throughput achieved by using the MT protocol. It can be concluded through the contents of this chapter that the DPM mechanism is capable of designing channel allocation protocols with complicated objectives in a multiuser multichannel CRN. In addition, the proposed queueing analytical framework was proved to be capable of providing comprehensive analysis of the performance metrics of a multi-user multi-channel CRN efficiently.

In Chapter five, a study with the help of the DPM mechanism to investigate the optimal performance that a multi-user multi-channel CRN is able to achieve was carried out. An optimization problem that is aimed at maximizing the overall weighted throughput of the entire system under the delay constraints was set up as a practical instance. The problem was formulated and solved under the framework of the DPM mechanism using a modified hill climbing algorithm proposed by the author. Techniques to simplify complex evaluation of the performance metric that costs unaffordable computational resources were presented in the chapter. The success of the optimization problem formulation and solution proves the ability of the DPM mechanism and the queueing analytical framework based on the DPM to explore the potential upper limit of overall performance and QoS that a multi-user multi-channel CRN is able to provide, given certain system settings and environmental conditions.

In conclusion, the DPM mechanism proposed in this thesis managed to fill in the gaps in the study of the RA protocol in CRN with the following contributions:

1. The protocols created by the DPM mechanism provides a flexible and adaptable pattern to assign multiple channels to uses in CRN.
2. The objectives of RA in CRN can be achieved by the protocol designed using the DPM mechanism.
3. The potential limit of overall performance of a CR system can be explored using the DPM mechanism.

6.2 RECOMMENDATIONS FOR FUTURE WORK

Although a lot of published work has been done to address the RA problem through the study of the allocation protocols, more work undoubtedly needs to be done to carry forward the research in this field. The following recommendations are based on the problems identified during the research reported in this thesis.

6.2.1 Recommendations on efficient algorithm design

The scale of the RA problem is related to the number of channels and number of users. Taking Chapter four for instance, the complexity of the algorithm to achieve global optimization grows exponentially with the increase in the number of channels and users. However, the computing time to arrange the RA occupies the limited and precious transmission time through the channels that are available to the CR system. The conflicts between the amount of time consumed by the calculation to chase better performance and the amount of time for the data transmission in the CRN can be eased, in the author's opinion, through the following research directions:

- Improve the implementation of the algorithm to obtain the solutions more efficiently. First, programing languages with inherent efficiency, such as the C language and Assembly language, can be used to implement the corresponding algorithms. Second, from the hardware perspective, parallel computing devices can be equipped in the CR allocator to further increase the speed of the calculations.
- Investigate the key factors that differentiate the users and channels in a multi-user multi-channel CRN, so that the users and channels that share identical features can be divided into groups, and the statistical behavior of the entire group, instead of every individual user or channel, can be used to describe the actions of each user or channel to reduce the complexity of the allocation algorithms.

6.2.2 Recommendations on consideration of practical factors

The realistic environment in which CR technology is applied is usually more complicated and volatile than the assumptions in research. Based on the consideration to modify the models to describe the real conditions more accurately, the components of the analytical framework of CRN can be improved to adapt the realistic situations. The following recommendations are made based on the study in this thesis.

- **Fading channel model**. It is common for CR users to have different characteristics that affects the fading pattern of the wireless channel, such as the velocity of the mobile station and obstructions between the transmitter and receiver. Thus, to make an accurate estimation of the data transmission capacity of each channel, the fading channel model of each user should fit its actual conditions. However, most existing research ignores the differences and applies identical fading channel models for all users. It is a promising research direction to develop a framework that can model the users using various fading patterns.

- **Model of user's demand**. Data transmission using CR technology is not guaranteed because any harmful interference with the licensed user must be avoided. Thus, suitable applications of CR technology may differ from the conventional communication system and the demands of such applications may vary according to the actual situations. In the author's opinion, data transmission with a high delay tolerance and low rate is an appropriate application of CR technology, such as regular non-urgent data exchange in wireless sensor networks. A specific model in light of the characteristics of CR technology to describe the user's demand can help improve and perfect the study on the application of CR technology.
- **Principles of RA**. Most of the research assumes that the resource allocation is simply controlled by a group of principles of the CR service provider. Considering the particularity of CR users, it is a good prospect to let the user participate in the process of planning the RA. Each user may evaluate the utility of the chance to utilize the provided resources and to obtain quotations for them. This change in the channel allocation principle can also maximize the utility of the CR service provider because the full use of the resources is made by assigning them to the user offering the highest bid.
- **Complex radio resource model**. The resource model in this research is additive, which means the sum of the resource shared by each users is fixed. However, in practice, the resource model is more complex. For example, in underlay CR or in the CR system applying non-orthogonal multiple access, the constraint condition of radio resources may be logarithmic or exponential and even time-variant. Research on how to build a model to describe such radio resource will move the CR research forward greatly and is worthy of careful study.

6.2.3 Recommendations on the standardization of the CR analytical framework

Numerous research programs are aimed at exploring every potential application of CR technology, since it has not been implemented extensively in reality. However, the models of the essential factors are built for their own purpose, and the supporting theory and ideas behind them vary. Thus, it is hard to compare ideas with different models and to determine what is the most appropriate solution in reality. A unified analytical framework for the CRN can serve as an objective tool to compare different research approaches aimed at the same goal. Methods to formulate the standardizations of such a framework are worth further study.

REFERENCES

[1] ITU. (2017) ITU world telecommunication/ICT indicators database. [Online]. Available: http://www.itu.int/en/ITU-D/Statistics/Pages/publications/wtid.aspx

[2] I. Peña-López *et al.*, "World development report 2016: Digital dividends," 2016.

[3] A. P. P. G. Steffes *et al.*, "Analysis and comparison of spectrum measurements performed in urban and rural areas to determine the total amount of spectrum usage," in *Proc. Int. Symp. Advanced Radio Technologies (ISART 2005)*, 2005, pp. 9–12.

[4] M. H. Islam, C. L. Koh, S. W. Oh, X. Qing, Y. Y. Lai, C. Wang, Y.-C. Liang, B. E. Toh, F. Chin, G. L. Tan *et al.*, "Spectrum survey in Singapore: Occupancy measurements and analyses," in *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*. IEEE, 2008, pp. 1–7.

[5] V. N. Q. Bao, T. D. Thuan, N. T. Quy, L. M. Trung *et al.*, "Vietnam spectrum occupancy measurements and analysis for cognitive radio applications," in *Advanced Technologies for Communications (ATC), 2011 International Conference on*. IEEE, 2011, pp. 135–143.

[6] Y. Liang, S.-x. YIN, W. Shuai, E.-q. ZHANG, W.-j. HONG, and S.-f. LI, "Quantitative spectrum occupancy evaluation in China: based on a large scale concurrent spectrum measurement," *The Journal of China Universities of Posts and Telecommunications*, vol. 19, no. 3, pp. 122–128, 2012.

[7] M. A. McHenry, D. McCloskey, and G. Lane-Roberts, "Spectrum occupancy measurements,

location 4 of 6: Republican national convention, New York City, New York, August 30, 2004-September 3, 2004, Revision 2," *Shared Spectrum Company Report*, 2005.

[8] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proceedings of the first international workshop on Technology and policy for accessing spectrum*. ACM, 2006, p. 1.

[9] S. D. Barnes, P. J. Van Vuuren, and B. Maharaj, "Spectrum occupancy investigation: measurements in South Africa," *Measurement*, vol. 46, no. 9, pp. 3098–3112, 2013.

[10] J. Mitola, "Cognitive radio—an integrated agent architecture for software defined radio," 2000.

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

[12] E. Biglieri, *Principles of cognitive radio*. Cambridge University Press, 2013.

[13] P. Marshall, *Quantitative analysis of cognitive radio and network performance*. Artech House, 2010.

[14] R. Yu, Y. Zhang, S. Gjessing, C. Yuen, S. Xie, and M. Guizani, "Cognitive radio based hierarchical communications infrastructure for smart grid," *IEEE network*, vol. 25, no. 5, 2011.

[15] A. Ghassemi, S. Bavarian, and L. Lampe, "Cognitive radio for smart grid communications," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE, 2010, pp. 297–302.

[16] R. C. Qiu, Z. Hu, Z. Chen, N. Guo, R. Ranganathan, S. Hou, and G. Zheng, "Cognitive radio network for the smart grid: Experimental system architecture, control algorithms, security, and microgrid testbed," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 724–740, 2011.

[17] O. Vermesan, P. Friess, P. Guillemin, S. Gusmeroli, H. Sundmaeker, A. Bassi, I. S. Jubert, M. Mazura, M. Harrison, M. Eisenhauer *et al.*, “Internet of things strategic research roadmap,” *Internet of Things-Global Technological and Societal Trends*, vol. 1, pp. 9–52, 2011.

[18] M. A. Shah, S. Zhang, and C. Maple, “Cognitive radio networks for Internet of Things: Applications, challenges and future,” in *Automation and Computing (ICAC), 2013 19th International Conference on*. IEEE, 2013, pp. 1–6.

[19] A. Aijaz and A. H. Aghvami, “Cognitive machine-to-machine communications for Internet-of-Things: A protocol stack perspective,” *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 103–112, 2015.

[20] H. Li and D. K. Irick, “Collaborative spectrum sensing in cognitive radio vehicular ad hoc networks: belief propagation on highway,” in *Vehicular technology conference (VTC 2010-spring), 2010 IEEE 71st*. IEEE, 2010, pp. 1–5.

[21] J. Wang, M. Ghosh, and K. Challapali, “Emerging cognitive radio applications: A survey,” *IEEE Communications Magazine*, vol. 49, no. 3, 2011.

[22] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, “Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds,” in *Internet of Things (WF-IoT), 2014 IEEE World Forum on*. IEEE, 2014, pp. 241–246.

[23] X. Hong, J. Wang, C.-X. Wang, and J. Shi, “Cognitive radio in 5g: a perspective on energy-spectral efficiency trade-off,” *IEEE Communications Magazine*, vol. 52, no. 7, pp. 46–53, 2014.

[24] *Standard for Wireless Regional Area Networks (WRAN) - Specific requirements - Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and procedures for operation in the TV Bands*, IEEE Std.

[25] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell, “IEEE 802.22: The first cognitive radio wireless regional area network standard,” *IEEE communications*

magazine, vol. 47, no. 1, pp. 130–138, 2009.

[26] M. Sherman, A. N. Mody, R. Martinez, C. Rodriguez, and R. Reddy, "IEEE standards supporting cognitive radio and networks, dynamic spectrum access, and coexistence," *IEEE Communications Magazine*, vol. 46, no. 7, 2008.

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

[28] L. Doyle, *Essentials of cognitive radio*. Cambridge University Press, 2009.

[29] B. Wang and K. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE Journal of selected topics in signal processing*, vol. 5, no. 1, pp. 5–23, 2011.

[30] B. A. Fette, *Cognitive radio technology*. Academic Press, 2009.

[31] B. Fette, "Fourteen years of cognitive radio development," in *Military Communications Conference, MILCOM 2013-2013 IEEE*. IEEE, 2013, pp. 1166–1175.

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

[33] S. Srinivasa and S. A. Jafar, "Cognitive radios for dynamic spectrum access-the throughput potential of cognitive radio: A theoretical perspective," *IEEE Communications Magazine*, vol. 45, no. 5, 2007.

[34] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications magazine*, vol. 46, no. 4, 2008.

[35] K.-C. Chen, Y.-J. Peng, N. Prasad, Y.-C. Liang, and S. Sun, "Cognitive radio network architecture: part I–general structure," in *Proceedings of the 2nd international conference on Ubiquitous information management and communication*. ACM, 2008, pp. 114–119.

[36] K.-C. Chen, Y.-J. Peng, N. Prasad, Y.-C. Liang, and S. Sun, "Cognitive radio network architecture: part II–trusted network layer structure," in *Proceedings of the 2nd international conference on Ubiquitous information management and communication*. ACM, 2008, pp. 120–124.

[37] D. Xu, Q. Zhang, Y. Liu, Y. Xu, and P. Zhang, "An architecture for cognitive radio networks with cognition, self-organization and reconfiguration capabilities," in *Vehicular Technology Conference (VTC Fall), 2012 IEEE*. IEEE, 2012, pp. 1–5.

[38] A. Amanna and J. H. Reed, "Survey of cognitive radio architectures," in *IEEE SoutheastCon 2010 (SoutheastCon), Proceedings of the*. IEEE, 2010, pp. 292–297.

[39] A. Goldsmith, S. A. Jafar, I. Marić, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, 2009.

[40] L. Kleinrock, *Queueing systems, volume 1: Theory*. wiley New York, 1975, vol. 66.

[41] L. Kleinrock, *Queueing systems, volume 2: Computer applications*. wiley New York, 1976, vol. 66.

[42] T. L. Saaty, *Elements of queueing theory: with applications*. McGraw-Hill New York, 1961.

[43] C. Newell, *Applications of queueing theory*. Springer Science & Business Media, 2013, vol. 4.

[44] A. S. Alfa, *Queueing theory for telecommunications: discrete time modelling of a single node system*. Springer Science & Business Media, 2010.

[45] J. N. Daigle, *Queueing theory with applications to packet telecommunication*. Springer Science & Business Media, 2005.

[46] T. G. Robertazzi, *Computer networks and systems: queueing theory and performance evaluation*. Springer Science & Business Media, 2012.

[47] J. Wang, A. Huang, L. Cai, and W. Wang, "On the queue dynamics of multiuser multichannel cognitive radio networks," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 3, pp. 1314–1328, 2013.

[48] M. M. Rashid, M. J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum scheduling for multiuser cognitive radio: a queueing analysis," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 10, pp. 5259–5269, 2009.

[49] A. J. Goldsmith and S.-G. Chua, "Adaptive coded modulation for fading channels," *Communications, IEEE Transactions on*, vol. 46, no. 5, pp. 595–602, 1998.

[50] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: the first worldwide wireless standard based on cognitive radios," in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*. Ieee, 2005, pp. 328–337.

[51] J. Khun-Jush, P. Schramm, U. Wachsmann, and F. Wenger, "Structure and performance of the HIPERLAN/2 physical layer," in *Vehicular Technology Conference, 1999. VTC 1999-Fall. IEEE VTS 50th*, vol. 5. IEEE, 1999, pp. 2667–2671.

[52] J. Khun-Jush, G. Malmgren, P. Schramm, and J. Torsner, "HIPERLAN type 2 for broadband wireless communication," *Europe*, vol. 5, pp. 5–35, 2000.

[53] T. ETSI, "101 475 V1. 3.1 (2001-12)," *Broadband radio access networks (BRAN)*, 2001.

[54] A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, "A comparison of the HIPERLAN/2 and IEEE 802.11 a wireless LAN standards," *Communications Magazine, IEEE*, vol. 40, no. 5, pp. 172–180, 2002.

[55] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels-a survey of principles and applications," *IEEE Signal Processing Magazine*, vol. 25, no. 5, 2008.

[56] C. E. Shannon, "Certain results in coding theory for noisy channels," *Information and control*, vol. 1, no. 1, pp. 6–25, 1957.

[57] D. Blackwell, L. Breiman, and A. Thomasian, "Proof of shannon's transmission theorem for finite-state indecomposable channels," *The Annals of Mathematical Statistics*, pp. 1209–1220, 1958.

[58] E. N. Gilbert, "Capacity of a Burst-Noise Channel," *Bell Labs Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[59] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[60] R. H. McCullough, "The binary regenerative channel," *Bell Labs Technical Journal*, vol. 47, no. 8, pp. 1713–1735, 1968.

[61] B. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE transactions on Information Theory*, vol. 13, no. 2, pp. 221–227, 1967.

[62] R. G. Gallager, *Information theory and reliable communication*. Springer, 1968, vol. 2.

[63] L. N. Kanal and A. Sastry, "Models for channels with memory and their applications to error control," *Proceedings of the IEEE*, vol. 66, no. 7, pp. 724–744, 1978.

[64] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Transactions on Information Theory*, vol. 35, no. 6, pp. 1277–1290, 1989.

[65] A. Semmar, M. Lecours, J.-Y. Chouinard, and J. Ahern, "Characterization of error sequences in UHF digital mobile radio channels," *IEEE transactions on vehicular technology*, vol. 40, no. 4, pp. 769–776, 1991.

[66] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Transactions on Communications*, vol. 39, no. 5, pp. 653–663, 1991.

[67] H. S. Wang and N. Moayeri, “Finite-state Markov channel - a useful model for radio communication channels,” *Vehicular Technology, IEEE Transactions on*, vol. 44, no. 1, pp. 163–171, 1995.

[68] C. C. Tan and N. C. Beaulieu, “On first-order markov modeling for the Rayleigh fading channel,” *IEEE Transactions on Communications*, vol. 48, no. 12, pp. 2032–2040, 2000.

[69] J. Aráuz and P. Krishnamurthy, “A study of different partitioning schemes in first order markovian models for Rayleigh fading channels,” in *Wireless Personal Multimedia Communications, 2002. The 5th International Symposium on*, vol. 1. IEEE, 2002, pp. 277–281.

[70] Q. Zhang, S. Kassam *et al.*, “Finite-state Markov model for Rayleigh fading channels,” *Communications, IEEE Transactions on*, vol. 47, no. 11, pp. 1688–1692, 1999.

[71] A. J. Goldsmith and P. P. Varaiya, “Capacity, mutual information, and coding for finite-state Markov channels,” *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 868–886, 1996.

[72] L. Li and A. J. Goldsmith, “Low-complexity maximum-likelihood detection of coded signals sent over finite-state Markov channels,” *IEEE transactions on communications*, vol. 50, no. 4, pp. 524–531, 2002.

[73] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, “Optimal and suboptimal packet scheduling over correlated time varying flat fading channels,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 446–456, 2006.

[74] H. Wang and N. B. Mandayam, “Opportunistic file transfer over a fading channel under energy and delay constraints,” *IEEE transactions on communications*, vol. 53, no. 4, pp. 632–644, 2005.

[75] F. Babich and G. Lombardi, “A Markov model for the mobile propagation channel,” *IEEE Transactions on Vehicular Technology*, vol. 49, no. 1, pp. 63–73, 2000.

[76] H. S. Wang and P.-C. Chang, “On verifying the first-order Markovian assumption for a Rayleigh fading channel model,” *IEEE Transactions on Vehicular Technology*, vol. 45, no. 2, pp. 353–357, 1996.

[77] B. Sklar, “Rayleigh fading channels in mobile digital communication systems. I. Characterization,” *IEEE Communications magazine*, vol. 35, no. 9, pp. 136–146, 1997.

[78] M.-S. Alouini and A. J. Goldsmith, “Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques,” *IEEE Transactions on Vehicular Technology*, vol. 48, no. 4, pp. 1165–1181, 1999.

[79] Y. R. Zheng and C. Xiao, “Simulation models with correct statistical properties for Rayleigh fading channels,” *IEEE Transactions on communications*, vol. 51, no. 6, pp. 920–928, 2003.

[80] W. Lindsey, “Error probabilities for rician fading multichannel reception of binary and N-ary signals,” *IEEE Transactions on Information Theory*, vol. 10, no. 4, pp. 339–350, 1964.

[81] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, “Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, 2006.

[82] K. A. Hamdi, “Capacity of MRC on correlated Rician fading channels,” *IEEE Transactions on Communications*, vol. 56, no. 5, pp. 708–711, 2008.

[83] M.-S. Alouini and A. J. Goldsmith, “Adaptive modulation over Nakagami fading channels,” *Wireless Personal Communications*, vol. 13, no. 1-2, pp. 119–143, 2000.

[84] N. C. Beaulieu and A. A. Abu-Dayya, “Analysis of equal gain diversity on Nakagami fading channels,” *IEEE Transactions on Communications*, vol. 39, no. 2, pp. 225–234, 1991.

[85] M.-S. Alouini, A. Abdi, and M. Kaveh, “Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels,” *IEEE Transactions on Vehicular Technology*, vol. 50, no. 6, pp. 1471–1480, 2001.

[86] F. Babich, G. Lombardi, and E. Valentinuzzi, "Indoor propagation characteristics in dect band," in *Proceedings of Vehicular Technology Conference - VTC*, vol. 1, Apr 1996, pp. 574–578 vol.1.

[87] M. Zorzi, R. R. Rao, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 2, pp. 445–455, 1997.

[88] G. Fairhurst and L. Wood, "Advice to link designers on link Automatic Repeat reQuest (arq)," 2002.

[89] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 5, pp. 1746–1755, 2004.

[90] L. L. Peterson and B. S. Davie, *Computer networks: a systems approach*. Elsevier, 2007.

[91] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 3, pp. 1142–1153, 2005.

[92] V. Srivastava and M. Motani, "Cross-layer design: a survey and the road ahead," *IEEE Communications Magazine*, vol. 43, no. 12, pp. 112–119, 2005.

[93] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications magazine*, vol. 41, no. 10, pp. 74–80, 2003.

[94] L. Georgiadis, M. J. Neely, L. Tassiulas *et al.*, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends® in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[95] *Error resilience in real-time packet multimedia payloads*, 3GPP TSG-S4 Codec Working Group, 1999.

[96] B. S. Awoyemi, B. T. Maharaj, and A. S. Alfa, "Solving resource allocation problems in cognitive radio networks: a survey," *EURASIP Journal on Wireless Communications and*

Networking, vol. 2016, no. 1, p. 176, 2016.

[97] B. Awoyemi, B. Maharaj, and A. Alfa, "Optimal resource allocation solutions for heterogeneous cognitive radio networks," *Digital Communications and Networks*, vol. 3, no. 2, pp. 129–139, 2017.

[98] B. S. Awoyemi, B. T. Maharaj, and A. S. Alfa, "Resource allocation in heterogeneous cooperative cognitive radio networks," *International Journal of Communication Systems*, vol. 30, no. 11, 2017.

[99] D.-h. Nam and H.-k. Min, "An energy-efficient clustering using a round-robin method in a wireless sensor network," in *Software Engineering Research, Management & Applications, 2007. SERA 2007. 5th ACIS International Conference on*. IEEE, 2007, pp. 54–60.

[100] M. Shreedhar and G. Varghese, "Efficient fair queuing using deficit round-robin," *IEEE/ACM Transactions on networking*, vol. 4, no. 3, pp. 375–385, 1996.

[101] E. L. Hahne, "Round-robin scheduling for max-min fairness in data networks," *IEEE Journal on Selected Areas in communications*, vol. 9, no. 7, pp. 1024–1039, 1991.

[102] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," *IEEE Journal on selected areas in communications*, vol. 26, no. 1, 2008.

[103] Y. Li and A. Nosratinia, "Hybrid opportunistic scheduling in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 328–337, 2012.

[104] M. M. Rashid, J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum access in cognitive radio networks: A queueing analytic model and admission controller design," in *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*. IEEE, 2007, pp. 4647–4652.

[105] Q. Chen, Y.-C. Liang, M. Motani, and W.-C. Wong, “A two-level MAC protocol strategy for opportunistic spectrum access in cognitive radio networks,” *IEEE Transactions on vehicular technology*, vol. 60, no. 5, pp. 2164–2180, 2011.

[106] L. Zhang, K. Zeng, and P. Mohapatra, “Opportunistic spectrum scheduling for mobile cognitive radio networks in white space,” in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. IEEE, 2011, pp. 844–849.

[107] R. Knopp and P. A. Humblet, “Information capacity and power control in single-cell multiuser communications,” in *Communications, 1995. ICC’95 Seattle,’Gateway to Globalization’, 1995 IEEE International Conference on*, vol. 1. IEEE, 1995, pp. 331–335.

[108] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, “Scheduling in a queuing system with asynchronously varying service rates,” *Probability in the Engineering and Informational Sciences*, vol. 18, no. 2, pp. 191–217, 2004.

[109] L. Zhang, Y. Xin, and Y.-C. Liang, “Weighted sum rate optimization for cognitive radio MIMO broadcast channels,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, 2009.

[110] S. Wang, B. T. Maharaj, and A. S. Alfa, “Resource allocation and performance measures in multi-user multi-channel cognitive radio networks,” in *Vehicular Technology Conference (VTC Spring), 2016 IEEE*.

[111] A. De Domenico, E. C. Strinati, and M.-G. Di Benedetto, “A survey on MAC strategies for cognitive radio networks,” *IEEE Communications Surveys & Tutorials*, vol. 14, no. 1, pp. 21–44, 2012.

[112] G. Yu, Z. Zhang, and P. Qiu, “Efficient ARQ protocols for exploiting cooperative relaying in wireless sensor networks,” *Computer Communications*, vol. 30, no. 14, pp. 2765–2773, 2007.

[113] M. Mitici, J. Goseling, J.-K. van Ommeren, M. de Graaf, and R. J. Boucherie, “On a tandem queue with batch service and its applications in wireless sensor networks,” *Queueing systems*, vol. 87, no. 1-2, pp. 81–93, 2017.

[114] J. G. Francis, “The QR transformation a unitary analogue to the LR transformation - Part 1,” *The Computer Journal*, vol. 4, no. 3, pp. 265–271, 1961.

[115] J. G. Francis, “The QR transformation - part 2,” *The Computer Journal*, vol. 4, no. 4, pp. 332–345, 1962.

[116] J. C. Nash, *Compact numerical methods for computers: linear algebra and function minimisation*. CRC press, 1990.

[117] K. Hamdi, W. Zhang, and K. B. Letaief, “Uplink scheduling with qos provisioning for cognitive radio systems,” in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*. IEEE, 2007, pp. 2592–2596.

[118] L. B. Le and E. Hossain, “Resource allocation for spectrum underlay in cognitive radio networks,” *IEEE Transactions on Wireless communications*, vol. 7, no. 12, pp. 5306–5315, 2008.

[119] Y. Wang, Y. Xu, L. Shen, C. Xu, and Y. Cheng, “Two-dimensional POMDP-based opportunistic spectrum access in time-varying environment with fading channels,” *Journal of Communications and Networks*, vol. 16, no. 2, pp. 217–226, 2014.

[120] L. B. Le, E. Hossain, and A. S. Alfa, “Delay statistics and throughput performance for multi-rate wireless networks under multiuser diversity,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3234–3243, 2006.

[121] J. Lee and H.-K. Park, “Channel prediction-based channel allocation scheme for multichannel cognitive radio networks,” *Journal of Communications and Networks*, vol. 16, no. 2, pp. 209–216, 2014.

[122] V. Rodriguez and D. J. Goodman, “Power and data rate assignment for maximal weighted throughput in 3G CDMA,” in *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, vol. 1. IEEE, 2003, pp. 525–531.

[123] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach (Third Edition)*. Prentice Hall PTR, 2009.

[124] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.