Some Cross-Layer Design and Performance Issues in Cognitive Radio Networks

by

S.M. Shahrear Tanzil

B.Sc., Bangladesh University of Engineering & Technology, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Electrical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

September 2013

Abstract

With the rapid deployment of data rate hungry wireless services, the demand for the radio spectrum is increasing day-by-day. On the other hand, fixed spectrum access (FSA) policy that allocates/assigns a certain portion of radio spectrum to a certain group of users has traditionally been adopted. Recent studies on spectrum usage have revealed that a large portion of the allocated spectrum is underutilized. In order to improve the overall spectrum utilization, recently dynamic spectrum access (DSA) policy has received a great deal of attention. Cognitive radio (CR) is a key enabling technology in order to facilitate DSA. CR can sense the radio spectrum and based on sensing outcome, it can adjust various transmission and operating parameters including bandwidth and power. Although CR technology in conjunction with DSA policy can improve the overall spectrum utilization, there are a number of challenges in designing CR based networks. Since the availability of radio spectrum for such networks is dynamic in nature, it is a quite challenging task to meet the quality of communications for the users in CR network (CRN). Another design challenge is spectrum sensing as an imperfect spectrum sensing can lead to a disturbance/collision to the original users (referred to as primary users) of the spectrum currently using the spectrum.

In order to address some of the above mentioned design challenges, in this thesis we make two major contributions, as follows. First, we develop resource allocation mechanisms that allocate available transmission rate of a particular CR user among its different classes of services using a cross-layer design approach that jointly considers the time varying nature of communication channels, availability of spectrum, and data link layer quality requirements of different classes of services.

In order to study the effect of imperfect sensing on data link layer's packet level performances as well as on collision probability, in the second part of this thesis, we develop a queuing analytic model that incorporates imperfect sensing. This analytic model is also useful for a call admission control decision in CRN when there is a certain sensing error as well as certain quality of service requirements for both primary and CR users. Using our developed model, we also compare performance of a random transmission protocol with that of the traditional deterministic transmission protocol.

Preface

I declare that I am the first author of this thesis. I developed all the analytic models and performed all the simulation results presented in this thesis. I also conducted literature survey on related topics and prepared drafts articles for scholarly publication. All these works were done under the supervision of Dr. Hossain. The work presented in Chapter 3 was done in collaboration with Dr. M Rashid who has suggested to use our developed model in Chapter 3 for a particular application that is call admission control decision.

A part of this thesis has been submitted for journal publication:

– S M Shahrear Tanzil, Md. Jahangir Hossain, and Mohammad M Rashid, "Rate allocation mechanisms for multi-class service transmission over cognitive radio networks", submitted in *IEEE Trans. on Wireless Commun.*, Sep. 2013.

Parts of this thesis have been accepted in two conferences:

– S M Shahrear Tanzil, Md. Jahangir Hossain, and Mohammad M Rashid, "Rate allocation mechanisms for multi-class service transmission over cognitive radio networks", accepted in *IEEE Global Commun. Conf. (GLOBECOM'13)*, Atlanta, GA, USA, Dec. 2013.

- S M Shahrear Tanzil and Md. Jahangir Hossain, "Cross-layer performance analysis for cognitive radio network with a random transmission protocol in presence of sensing errors", in Proc. of the *Int. Conf. on Cognitive Radio Oriented Wireless Networks (CROWNCOM'13)*, Washington DC, USA, Jul. 2013.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ARQ	Automatic repeat request
BE	Best effort
BS	Base station
CAC	Call admission control
CR	Cognitive radio
CRN	Cognitive radio network
DS	Delay sensitive
DSA	Dynamic spectrum access
FSA	Fixed spectrum access
LP	Linear programming
MAC	Medium access control
MDP	Markov decision process
OFDM	Orthogonal frequency division multiplexing
PHY	Physical
PU	Primary user
QBD	Quasi-birth-death
QoS	Quality of service

SU	Secondary user

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Dr. Jahangir Hossain for his guidance and persistent help throughout the research. This thesis would not have been possible without his support. I am greatly indebted to Dr. Hossain for his continuous encouragement and valuable suggestions.

I would like to thank all the members of my research group for their encouragement. Last, but not least, I would like to express my gratefulness to my family.

Chapter 1

Introduction

Wireless communication has become an integral part of our life and due to the technological advances day-by-day the number of wireless devices and services are increasing at an enormous rate. These wireless devices and services can require higher data rates which in turn require higher bandwidth. However, suitable portion of the radio spectrum is limited. Therefore, with the rapid deployment of new wireless services, the demand for radio spectrum is increasing day-by-day [1]. On the other hand, fixed spectrum access (FSA) policy has traditionally been adopted [2]. According to the FSA policy a certain portion of radio spectrum is allocated/reserved for a certain group of users by the spectrum regulatory bodies. In a given geographical location, only these users can access their assigned portion of the radio spectrum on a long term basis. In other words, these users have exclusive rights to use the allocated spectrum at any time in a given geographical location. These users to whom the spectrum is originally assigned are usually referred to as primary users (PUs) in the literature. With the FSA policy, other group of potential users, usually referred to as secondary users (SUs) are not allowed to access the spectrum, even if a particular portion of the spectrum is currently not being used by the PUs. Recent studies on spectrum measurements have revealed that a large portion of the assigned spectrum is used sporadically by the PUs [3] however, it is difficult to find a suitable portion of spectrum that can be given access or allocated to a new wireless service. Field measurements and studies indicate that FSA policy leads to an inefficient usage of valuable spectrum resource. For example according to the Federal Communications Commission (FCC), only 5.2% of the spectrum is utilized in United States below 3 GHz in a given location in a given time [4]. Therefore, spectrum scarcity is apparently an artificial problem and this observation has motivated researchers as well as various spectrum regulatory bodies around the world to find a solution for efficient utilization of precious spectrum resource. In order to improve overall spectrum utilization, dynamic spectrum access (DSA) policy has been considered as an alternate spectrum access policy. For example, a standard is being developed by IEEE known as IEEE802.22 in order to allow sharing of unused spectrum allocated to television broadcast service in a given geographical location to provide broadband access to the rural population. DSA policy is described below in details.

1.1 Dynamic Spectrum Access

According to the DSA policy, a group of potential SUs to whom the spectrum is not allocated/assigned can share the spectrum with the PUs opportunistically. In particular, at a given time in a given geographical location, the SUs can use a particular portion of the spectrum if that particular portion of the spectrum is not used by the PUs. With DSA policy, the spectrum can be accessed by SUs in two different ways namely, *spectrum underlay* and *spectrum overlay* [5] as described below.

1.1.1 Spectrum Underlay

In a spectrum underlay method, a SU can simultaneously transmit at the same spectrum with a PU provided that the interference caused by the SU is kept within a certain threshold which is termed as interference temperature limit [6]. With this model, in order to maintain the interference threshold limit prescribed by the PUs' networks or regulatory bodies, the SUs' network has to control the transmit power [7]. Based on the interference limit, the transmit power from a secondary network's transmitter can be low. As such the coverage of the SUs' network can be small. Other limitations of spectrum underlay access mechanisms include a higher interference from a primary transmitter due to the co-existence of both PUs and SUs in the same spectrum at the same time.

1.1.2 Spectrum Overlay

In a spectrum overlay method, in a given geographical location in a given time SUs' system senses the spectrum of the PUs and identifies the portion of the spectrum that are not being used by the PUs. The vacant portion of the spectrum is also termed as spectrum hole or slot. The SUs utilize the spectrum holes for their transmission until the PUs arrive back to use the portion of the spectrum. When a PU arrives and starts to use the spectrum, a SU has to evacuate the spectrum band for PUs' usage. A spectrum overlay method is illustrated in Fig. 1.1. With this model, the SUs' transmitter can transmit relatively higher amount of power and their network coverage can be high. However, the availability of radio spectrum for SUs' communication is dynamic in nature that depends on the activity of the PUs in a particular spectrum band. Also the SUs' network requires reliable sensing mechanisms to identify spectrum holes or slots. A sensing error can lead to a collision of SUs transmission with that of the PUs. Various sensing mechanisms have been proposed in the literature.

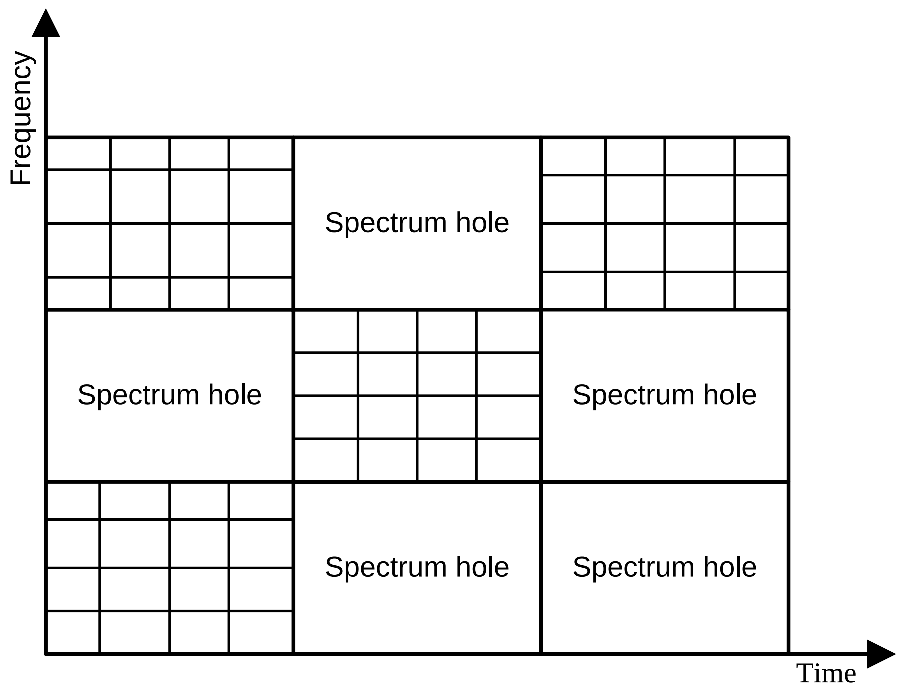

Figure 1.1: An example of overlay spectrum access with spectrum holes [2].

1.2 Cognitive Radio Design Philosophy and Cognitive Radio Network

Cognitive radio (CR) is an innovative radio design philosophy that can facilitate DSA. According to this radio design philosophy, the radio transmitters can adjust various transmission and operating parameters including the frequency range, modulation type, and power according to the wireless environment [8]. The concept of CR was first proposed by Joseph Mitola III in 1998. Mitola described CR as [9]: *The point in which wireless personal digital assistants (PDAs) and the related networks are sufficiently computationally intelligent about radio resources and related computer-to-computer communications to detect user communications needs as a function of use context, and to provide radio resources and wireless services most appropriate to those needs.* FCC defined CR as [10]: *A radio that can change its transmitter parameters based on interaction with the environment in which it operates.* In general, CR refers to a radio system that can sense the spectrum and based on the sensing outcome, it adapts the operating and transmission parameters [11]. It differs from the conventional radio with enhanced features such as cognition capability and reconfigurability [12], [13]. Identifying available spectrum and gathering information of the environment refers to the cognition capability. Reconfigurability refers to the adaptability of parameters according to the environment in order to use the radio spectrum efficiently. Since SUs access radio spectrum opportunistically, they are also referred to as SUs in the literature and to be consistent throughout the rest of this thesis we use SUs to refer those users that opportunistically access spectrum of PUs.

A cognitive radio network (CRN) consists of a set of SUs with or without an infrastructure e.g., a base station (BS). Such network uses PUs' spectrum as it does not have any allocated/licensed spectrum. Different architectures of cognitive networks are envisioned and studied in the literature. According to the network architecture, CRN can mainly be classified as an infrastructure-based CRN and ad-hoc CRN [14]. In this thesis, we consider an infrastructure-based CRN and a detailed description of such networks is given in Chapter 2.

1.3 Literature Review and Motivation

Although CR technology can improve the overall spectrum utilization, there is a number of challenges in designing CRN. One of the design challenges is the dynamic availability of radio spectrum for communications of such networks. Different physical (PHY) layer performance metrics e.g., achievable capacity, throughput of CR system for such dynamic environment were analyzed in [15], [16], [17]. Various network layer and data link layer level performance metrics were also analyzed for such networks. In particular, via a cross-layer approach, various packet level performances at the data link layer buffer e.g., packet queuing delay and packet loss probability of CRN were analyzed in [18] considering the dynamic availability of radio spectrum. In [19], [20], a G/M/N queuing model was used to analyze the call-level performance of the CRN. In [21], a M/G/1 queuing system was considered and the performance of CRN was evaluated for different traffic and channel conditions. In [22], a M/G/1-based queuing model was developed to characterize the performance of multi-channel MAC protocols for CRN. In order order to cope with the dynamic availability of radio spectrum, innovative medium access (MAC) layer protocols have been developed in the literature. In particular, various channel allocation mechanisms are developed in the literature see for examples, [6], [23] and [24]. All these works considered that SUs in the network to support only a single class of traffic.

The main motivation of the first part of this thesis came from the fact that now-a-days due to the advancement on wireless devices, most of the wireless systems support multi-class services e.g., video conferencing, email transfer and web browsing over wireless networks [25]. One can easily project that the future wireless systems based on CR technology will also be required to support these services. Different services have different traffic characteristics as well as different quality of service (QoS) e.g., delay, packet loss probability and throughput requirements [26], [27]. In terms of delay requirements, services can be classified as delay sensitive (DS) service or delay non-sensitive service [28]. Delay non-sensitive service also known as best effort (BE) service has no stringent delay requirements. One example of BE service is email service and one example of DS service is video conferencing. So, one important question is how to develop innovative resource e.g., transmission rate allocation mechanisms that can meet diverse QoS requirements of different classes of services transmitted over the CRN where not only the channel quality varies but also the availability of radio spectrum is very dynamic [29]. Although some efficient resource allocation mechanisms [30], [31], [32] and QoS aware call admission controller (CAC)[1] [33], [34] have been devised in the literature, they focused on only single class of traffic.

Another design challenge in implementing CRN is the reliability of spectrum sensing mechanisms. Several PHY layer channel sensing/detection methods e.g., energy detection, waveform detection, cyclostationarity detection and matched filter detection have been proposed in the literature [35]. Most of the sensing mechanisms can have a certain probability of sensing error [36], [37]. In general, there are two types of sensing errors as follows. CRN may not be able to detect an active PU in a given channel. Such sensing error is referred to as a miss-detection. On the other hand, CRN may detect a channel being used by a PU where in reality the channel is idle. This type of sensing error is known as a false alarm. Due to a miss-detection, transmission of CRN leads to a collision with the transmission of PUs. Therefore, sensing errors affect the overall performance of both PUs and SUs [38] and it is important to study the effect of imperfect sensing errors on data link layer's packet level performances as well as on collision probability.

In order to address some of the above mentioned design challenges in

[1]Call admission control (CAC) refers to a process in communication system that determines whether a new connection from a user should be accepted or rejected. This process checks out current resources and determines whether a new connection can be accommodated or not while satisfying all QoS requirements of all active connections in the system.

CRN, in this thesis we make two major contributions, as follows.

1.4 Contributions

- In this thesis, we study multi-class service transmission over CRN. In particular, we develop and study resource allocation mechanisms that allocate available transmission rate of a particular SU among its two different classes of services using a cross-layer design approach that jointly considers the time varying nature of communication channels, availability of spectrum, and data layer quality requirements of different classes of services. We formulate the rate allocation mechanism of a SU between its two different classes of services namely, DS and BE services as a Markov decision process (MDP). Then the optimal rate allocation mechanism that minimizes the average queuing delay of DS service while guaranteeing the packet loss probabilities of both classes of services is obtained using a linear programming (LP) technique. Since the optimal rate allocation mechanism can be complex to implement in practice, we propose and study a low-complexity suboptimal rate allocation mechanism. For this suboptimal scheme, we develop a queuing analytic model in order to measure different quality parameters e.g., packet loss probability, average queuing delay and queuing delay distribution. The analytic model is useful for a CAC decision in CRN when the suboptimal rate allocation mechanism is employed.

- In oder to study the effect sensing errors on different performance parameters e.g., collision probability, packet loss probability and queuing delay, in the second part of this thesis, we develop a queuing analytic model that incorporates imperfect sensing. This analytic model is also useful for CAC decision in CRN when there is a certain sensing error as well as certain quality of service requirements for both primary and SUs. Using our developed model, we also compare the performance of a random transmission protocol with that of the traditional deterministic transmission protocol.

1.5 Thesis Outline

The remainder of the thesis is organized as follows. In Chapter 2, we describe some basic background information on infrastructure-based CRN, quasi-birth-death (QBD) process and MDP. In Chapter 3, we develop optimal and suboptimal rate allocation mechanisms for multi-class service transmission over CRN and study their performances. In oder to study the effect of sensing errors on collision probability, packet delay, and loss probabilities, in Chapter 4, we develop a queuing analytic model that incorporates sensing errors, time varying nature of channels, as well as dynamic availability of the channels. In order to illustrate the usefulness of the developed model, we also provide an example in this chapter.

Chapter 2

Background

This chapter presents some background information on infrastructure-based CRN along with the operating assumptions, QBD processes and discrete time MDP.

2.1 Infrastructure-based Cognitive Radio Network

2.1.1 Overall Description

As mentioned in Chapter 1, according to the network architecture, CRN can be classified as either infrastructure-based CRN or ad-hoc CR network. In this thesis, we consider an infrastructure-based CRN. The main advantage of the infrastructure-based CRN is that it provides a greater coverage area. In what follows, we provide a general description and describe operating assumptions of an infrastructure-based CRN that is under consideration. A typical infrastructure-based CRN is shown in Fig. 2.1 where a CR BS supports K different SUs. The CRN opportunistically uses a total bandwidth of B_{W} Hz which is divided into one or more orthogonal channel(s). The bandwidth is licensed/allocated to a primary system who has the priority

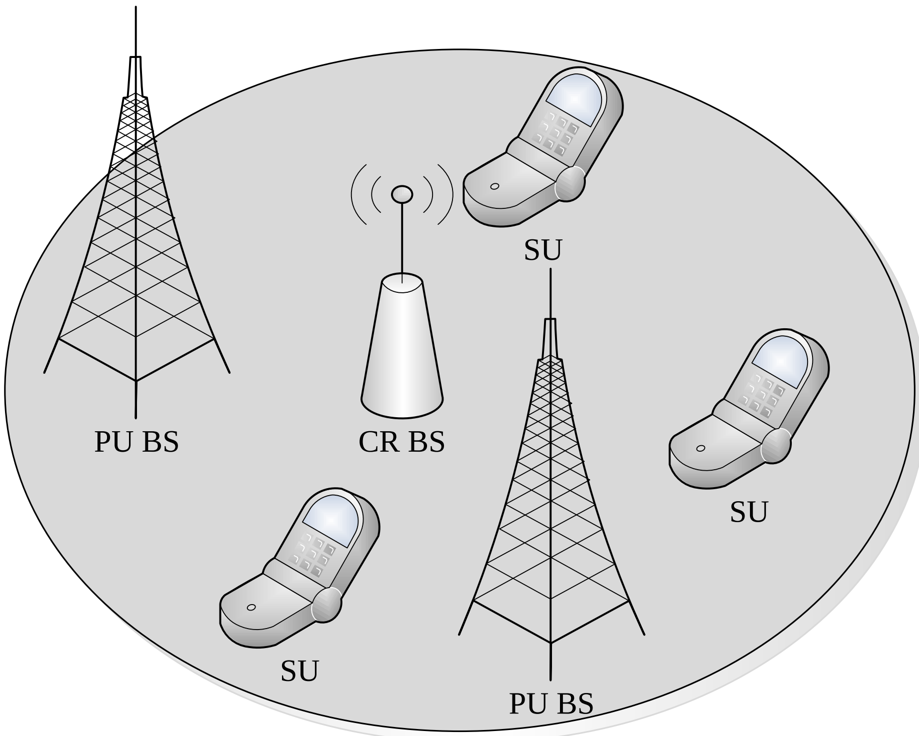

Figure 2.1: Infrastructure-based cognitive radio network.

right for using it. It is assumed that both primary system and CRN work in a slotted fashion with a fixed slot duration of T_f sec. CRN shares the spectrum with the PUs using overlay spectrum access method. In particular, at the beginning of a time slot, CRN senses the channel(s) in order to estimate whether the channel(s) are occupied or not. In a given time slot, CRN uses the channel(s) opportunistically. It is assumed that in a given transmission slot the channel can be assigned to only one SU using so-called max rate channel scheduling algorithm. A detailed description of the max rate channel allocation scheme is provided in later section.

Packet(s) arrive randomly at the data link layer from the upper layer of communication protocol stack. Those packets are temporarily stored in the data link layer's buffer to be transmitted over the wireless channel. It is assumed that the channel quality varies over time and transmission rate is adapted according to the channel quality. In what follows we provide brief descriptions of PUs' activity, channel fading and rate adaptation, max rate channel scheduling, and random packet arrival process.

2.1.2 Primary Users' Activity

In order to model the behavior or activity of a PU in a particular channel, ON-OFF model is commonly used in the literature [39], [40], [41], [42]. According to the ON-OFF model, PUs' activity on a particular channel can be modelled by a two state time homogeneous first order Markov process as shown in Fig. 2.2.

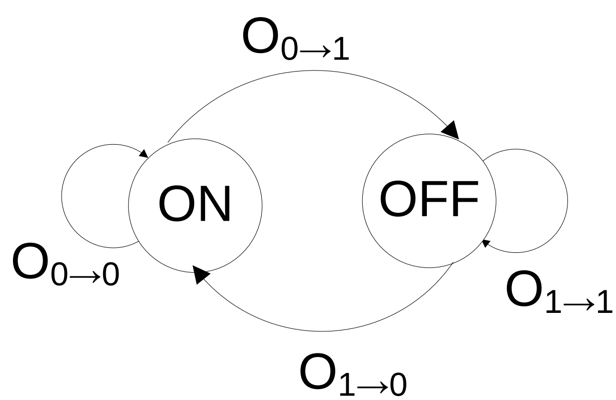

Figure 2.2: An ON-OFF model for primary users' activity in a particular channel.

PUs' channel occupancy state for a particular channel is denoted by O where $O \in \{0, 1\}$. If the channel is occupied by a PU i.e., busy/ON, O takes value 0. Otherwise, O takes value 1. The transition probability matrix of the channel occupancy state is defined as follows

$$\mathbf{P}_{\mathrm{O}} = \begin{bmatrix} O_{0\to 0} & O_{0\to 1} \\ O_{1\to 0} & O_{1\to 1} \end{bmatrix}, \tag{2.1}$$

where $O_{l\to j}$ denotes the transition probability from state l to state j and $l, j \in \{0, 1\}$. PUs' activity, ρ in a channel is defined as the percentage of time on an average the channel is occupied by PUs. This corresponds to the steady state probability of a particular channel being occupied by PUs and can be written in a closed-form as follows

$$\rho = \frac{1 - O_{1\to 1}}{2 - O_{0\to 0} - O_{1\to 1}}. \tag{2.2}$$

On an average if the channel remains occupied for a longer duration by the PUs, the value of ρ increases and vice-versa.

2.1.3 Channel Fading Model and Rate Adaptation

The channel quality of a particular channel between the CR BS and a SU is considered as a slowly time varying channel i.e., the channel fading amplitude remains roughly constant over the slot duration, T_f [18], [43], [44] however it changes in the next time slot. It is assumed that the channel fading amplitude follows Nakagami-m distribution which represents a range of practical fading distributions via the m parameter [18]. Therefore, the received signal to noise ratio (SNR), ϕ follows the gamma distribution. Due to the slowly time varying nature of the channel, the fading process can be modelled as a finite state Markov channel (FSMC) [18], [45] where the possible channel fading states are denoted by a set $\mathcal{C} = \{0, 1, ..., Z-1\}$ with total Z states. At a particular time slot, the channel is in state z if $A_z \leq \phi \leq A_{z+1}$ where A_z is the state boundary [45]. According to the FSMC model, it is also assumed that SNR, ϕ remains in the same state over a time slot and transition can occur only in the same state or adjacent states at the next time slot. One can obtain the channel state transition probability from state z to state $z+1$, $\Pr(C_{z,z+1})$ as follows [46], [47]

$$\Pr(C_{z,z+1}) = \frac{N_{R_{z+1}}}{\Pr(C_z) \times \text{transmission rate}}, \tag{2.3}$$

where $N_{R_{z+1}}$ represents the expected number of times per second the received SNR, ϕ crosses the state boundary, A_{z+1} in a downward direction and $\Pr(C_z)$ represents the steady state probability of channel being in state z. $N_{R_{z+1}}$ and $\Pr(C_z)$ can be calculated using the procedure mentioned in [47].

Adaptive modulation and coding (AMC) is employed by CRN at the transmitter in order to exploit the time varying nature of the wireless channels [48]. In particular, in a given time slot, modulation order and/or coding rate are adjusted according to the channel fading state. Therefore, for a fixed packet size, the number of packets that can be transmitted during a transmission slot can vary. If the channel is in state z ($z \in \mathcal{C}$) at a particular time slot, rz packet(s) can be transmitted through the channel where r is an integer depends on slot duration, modulation and coding parameters.

2.1.4 Max Rate Channel Scheduling

Since the CR BS supports multiple SUs, a channel can be allocated to only one SU in a given time slot. In order to allocate the channel among active SUs, so-called max rate channel scheduling mechanism is considered as it can maximize the overall throughput of CRN [18], [49]. According to the max rate channel scheduling scheme, the CRN assigns a channel to a SU who can support the highest transmission rate in that particular channel. If more than one SU can support the highest transmission rate in that particular channel, CRN assigns the channel to a SU randomly with equal probability among these SUs.

2.1.5 Packet Arrival Process

In general, packet arrival from the upper layer of communication protocol stack at the data link layer is random in nature. In order to model this random arrival, Poisson process, Markov modulated Poisson process, and batch Bernoulli process are commonly used [50], [51], [18]. In this thesis, we consider batch Bernoulli random process for packet arrival at the data link layer buffer as it can capture different burstiness in traffic arrival process [18]. In general, batch Bernoulli model can be described by $\boldsymbol{\beta} = [\beta_0, \beta_1, ..., \beta_N]$, where N is the maximum number of packets that can arrive at a time slot and β_j $(0 \leq j \leq N)$ denotes the probability of j packet(s) arrival at a particular time slot. Average packet arrival rate at the data link layer can be calculated as follows

$$\beta' = \sum_{j=0}^{N} j\beta_j. \tag{2.4}$$

2.2 Quasi-Birth-Death Process

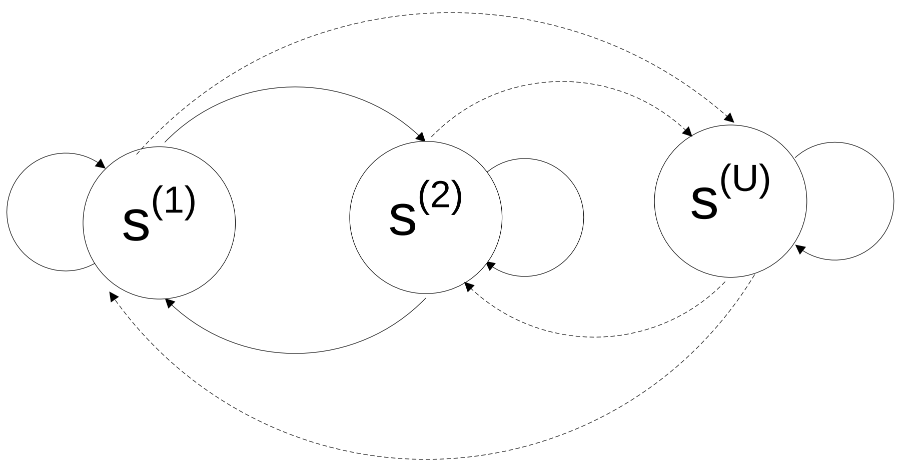

Figure 2.3: An example of Markov chain.

A birth-death (BD) process is a special case of a Markov chain (MC) where MC is a random process that transits from one state to another state randomly. An example of a MC with total $s^{(U)}$ states is shown in Fig. 2.3. In a BD process, transitions are restricted to only adjacent or same states as shown in Fig. 2.4. A QBD process is a special case of a BD process as described below.

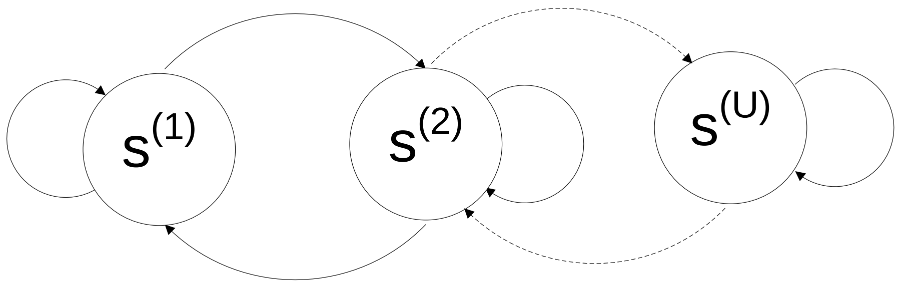

Figure 2.4: An example of a birth-death process.

Consider a MC with two dimensional state space, $\{(i,j)|0 \leq i, 0 \leq j \leq U_i\}$ where i represents the level and j represents the state. Level l consists of a set of states, $\{(l,0),(l,1),...,(l,Ul)\}$. Such a MC is called a QBD process if transitions are restricted to its nearest neighbor levels [52], [53]. An example of a QBD process is shown in Fig. 2.5 where there are three levels namely level 0, level 1 and level 2. Levels 0, 1 and 2 have $U0$, $U1$ and $U2$ states, respectively. The QBD process is defined by its transition probability matrix. For an example, the transition probability matrix of a QBD process with an infinite number of states is written below

$$\mathbf{P}_{\text{inf}} = \begin{bmatrix} \mathbf{A} & \mathbf{C} & & & & \\ \mathbf{D} & \mathbf{E} & \mathbf{F_0} & & & \\ & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & \\ & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & \\ & & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{2.5}$$

where $\mathbf{A}$ and $\mathbf{C}$ represent the transition probability sub-matrices of level 0, $\mathbf{D}$, $\mathbf{E}$ and $\mathbf{F_0}$ represent the transition probability sub-matrices of level 1 and

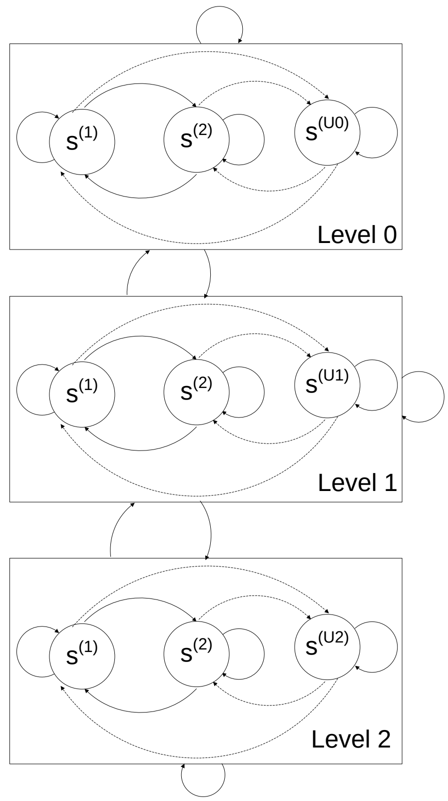

Figure 2.5: An example of a quasi-birth-death process.

so forth [54]. For a finite state space, we can rewrite equation (2.5) with level X as follows

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ X-1 \\ X \end{array} \begin{bmatrix} \mathbf{A} & \mathbf{C} & & & & & \\ \mathbf{D} & \mathbf{E} & \mathbf{F_0} & & & & \\ & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & & \\ & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F'_0} \\ & & & & & \mathbf{F'_2} & \mathbf{F'_1} \end{bmatrix}. \tag{2.6}$$

Level 0 and X are called boundary levels, level 1 and $X-1$ are called border levels and level 2, 3,....,$X-2$ are called repeating levels. From eq. (2.6) one can obtain generator matrix, $\mathbf{G}$ as follows [54]

$$\mathbf{G} = \mathbf{P} - \mathbf{I} \tag{2.7}$$

where $\mathbf{I}$ is an identity matrix. For an example the generator matrix corresponding to the transition matrix $\mathbf{P}$ in eq. (2.6) can be written as

$$\mathbf{G} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ X-1 \\ X \end{array} \begin{bmatrix} \mathbf{A_G} & \mathbf{C_G} & & & & & \\ \mathbf{D_G} & \mathbf{E_G} & \mathbf{F_{0G}} & & & & \\ & \mathbf{F_{2G}} & \mathbf{F_{1G}} & \mathbf{F_{0G}} & & & \\ & & \mathbf{F_{2G}} & \mathbf{F_{1G}} & \mathbf{F_{0G}} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \mathbf{F_{2G}} & \mathbf{F_{1G}} & \mathbf{F'_{0G}} \\ & & & & & \mathbf{F'_{2G}} & \mathbf{F'_{1G}} \end{bmatrix}, \tag{2.8}$$

where $\mathbf{A_G}$ and $\mathbf{C_G}$ represent the sub-matrices of level 0 of generator matrix and so forth [54]. Matrix-geometric is a well-known method to find the steady state probabilities of a QBD process [54]. Let us use $\mathbf{v}_l$ to denote the vector of steady state probabilities of states in level l. The detailed procedure to calculate steady state probabilities is mentioned in the Appendix D.

2.3 Markov Decision Process

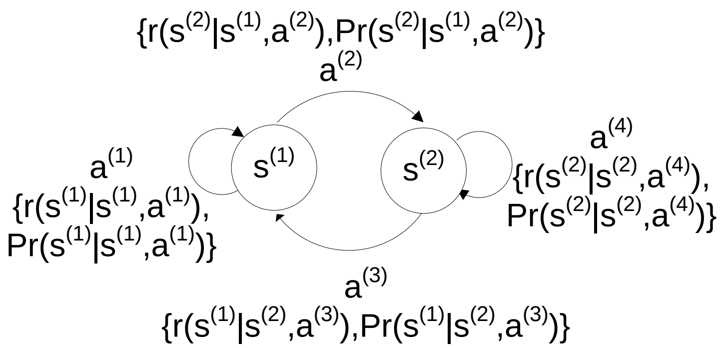

Figure 2.6: An example of a two state Markov decision process [55].

A discrete time MDP is a stochastic control process and provides a mathematical framework to model sequential decision making under uncertainty [55]. A discrete time MDP is described through a set of time slots $\mathcal{T} = 1, 2, .., N_t$, a set of states $\mathcal{S} = \{s^{(1)}, s^{(2)}, ..., s^{(U)}\}$, a set of actions $\mathcal{A} = \{a^{(1)}, a^{(2)}, ..., a^{(Y)}\}$, a set of immediate costs that depend on states and actions and a set of state and action dependent transition probabilities. An example of a two state MDP is shown in Fig. 2.6. Without loss of generality, let us assume that the process is in state $s^{(1)}$ at time slot n where $n \in \mathcal{T}$. The decision maker or controller chooses action $a^{(2)}$ from the set of possible actions in state $s^{(1)}$, $\mathcal{A}_{s^{(1)}} = \{a^{(1)}, a^{(2)}\}$ and at time slot $n+1$, the process moves to the state $s^{(2)}$ with transition probability $\Pr(s^{(2)}|s^{(1)}, a^{(2)})$. For taking action $a^{(2)}$, the decision maker receives an immediate expected reward/cost, $r(s^{(2)}|s^{(1)}, a^{(2)})$. At time slot $n+1$, the process is in state $s^{(2)}$. The decision maker chooses action $a^{(3)}$ where $a^{(3)} \in \mathcal{A}_{s^{(2)}}$ and $\mathcal{A}_{s^{(2)}} = \{a^{(3)}, a^{(4)}\}$ denotes the feasible set of actions at state $s^{(2)}$. The process moves to the state $s^{(1)}$ with transition probability $\Pr(s^{(1)}|s^{(2)}, a^{(3)})$ and the decision maker receives an immediate expected reward/cost, $r(s^{(1)}|s^{(2)}, a^{(3)})$. The decision maker takes a sequence of actions/decisions and receives a sequence of rewards/costs as the process moves forward. The outcome of the MDP is partly under the control of the decision maker via taking a particular action at a particular state and partly random due to the random transition probabilities [56]. The objective of the decision maker is to select a set of actions that maximizes his/her overall/cumulative reward or minimizes the overall/cumulative cost.

MDP has extensively been used to formulate optimization problems in many areas such as operation research, management science and communication network engineering [56]. In general, wireless network optimization problems sometime require sequence of decisions while the system moves from one state to another and there exist trade-offs between different rewards/costs e.g., delay, packet loss, energy [56]. Such sequential decision making problems can be formulated as a constrained MDP where one type of cost is minimized while keeping other costs below target threshold values

[57]. Then the optimal sequence of actions for such constrained MDPs can be solved via a LP or a combination of the Lagrangian multiplier method [56]. The detailed procedure of formulating a constrained MDP in the context of rate of allocation mechanism for multi-class data transmission over CRN is mentioned in Chapter 3.

Chapter 3

Multi-Class Service Transmission over Cognitive Radio Networks

[2]As described in Chapter 1, supporting multi-class services over CRN is quite challenging due to dynamic availability of radio spectrum, the time varying nature of the channel as well as diverse QoS requirements of multi-class services. Various network layer and data link layer level performance metrics were also analyzed for such networks [18], [19], [20], [21], [22]. A M/G/1-based queuing model was developed to characterize the performance of multi-channel MAC protocols for CRN. In order order to cope with the dynamic availability of radio spectrum, innovative medium access (MAC) layer protocols have been developed in the literature. In particular, various channel allocation mechanisms are developed in the literature see for examples, [6], [23] and [24]. All these works considered that SUs in the network

[2]The research work presented in this chapter has been accepted in part as S M Shahrear Tanzil, Md. Jahangir Hossain, and Mohammad M Rashid, "Rate allocation mechanisms for multi-class service transmission over cognitive radio networks", *IEEE Global Commun. Conf. (GLOBECOM'13)*, Atlanta, GA, USA, Dec. 2013. A part of this work is also submitted in *IEEE transaction on wireless communication.*

to support only a single class of traffic.

Recently multi-class service transmission over CRN has recently received a great deal of attention [58], [59], [60], [61]. In [58], [59] and [61], authors considered a CRN that supports two types of SUs where SUs are prioritized based on their services i.e., SUs supporting DS services are given higher priority over SUs supporting BE services. In particular, authors of [58] developed a call-level spectrum hand-off mechanism based on the priorities of SUs. They also investigated the impact of channel reservation scheme for higher priority SUs. Although the channel reservation scheme can improve the QoS for high priority SUs, it leads to a higher blocking probability for the low priority SUs. Therefore, in [59] authors introduced a queue where service requests from low priority SUs can be queued to be served later on when the channels are available for their communications. In [60] authors prioritized active SUs in the system according to the bit error rate, delay and throughput requirements and developed a spectrum allocation mechanism among SUs based on these priorities. Authors of [61] proposed a queuing system with two different queues where SUs with different priorities can be queued. Then available channels can be allocated proportionately among these SUs based on the priorities of the SUs. These works considered prioritizing SUs at the connection/user level and analyzed/improved various connection level performance metrics e.g., call blocking, dropping and interruption probabilities.

Unlike the above works, in this chapter, we consider a system model where each SU in CRN simultaneously supports two different classes of services namely BE service (i.e., delay non-sensitive) and DS service and study resource allocation mechanisms between these services. In particular, we formulate a rate allocation mechanism of a SU between its two different classes of services as a MDP. For this formulation, the buffer dynamics are expressed in a generic form. Then the optimal rate allocation mechanism that minimizes the average queuing delay of DS packets while guaranteeing packet loss probabilities of both classes of services is obtained using a LP technique. Since, the optimal rate allocation mechanism can be complex to implement in practice, we propose a low-complexity suboptimal rate allocation mechanism and study its performance. In particular, for this suboptimal scheme, we develop a queuing analytic model which is analyzed as a QBD process in order to measure different packet-level performance parameters. Selected numerical results show that the performance of the suboptimal rate allocation mechanism is quite similar to the optimal rate allocation mechanism. Our developed queuing analytic model with the suboptimal rate allocation mechanism is useful not only to measure packet-level performance parameters but also to design a CAC for a CRN that supports SUs each with two different classes of services.

The rest of the chapter is organized as follows. Section 3.1 describes the system model whereas Section 3.2 presents formulation of the optimal rate allocation mechanism. In Section 3.2, we also present the suboptimal rate allocation mechanism along with the developed queuing analytic model to measure various performance parameters. In Section 3.3, we present some numerical results and in Section 3.4 we demonstrate a possible application of our analytic model.

3.1 System Model

3.1.1 Overall Description

As mentioned in Chapter 2, an infrastructure-based CRN is considered in this thesis. We also consider that total bandwidth, B_{W} is divided into N_p channels. In a given time slot, the CRN senses all the channels at the beginning. We assume that the channels can be sensed perfectly i.e., the channel sensed as idle is actually idle and channel sensed as busy is actually busy. From the channel sensing, CRN identifies the channels that are not occupied by the PUs in that particular time slot. Then CRN assigns these available empty channels among its SUs using the so-called max rate channel scheduling mechanism as described in Section 2.1.4.

We consider that each SU simultaneously supports two different classes of services namely BE service and DS service. Correspondingly, there are two data link layer buffers at each SU's terminal as shown in Fig. 3.1 where only kth SU's buffers are shown. Packets arriving from the upper layer are stored in these data link layer's buffers before they are transmitted over the wireless channels. In our system model, uplink i.e., SUs to CR BS transmission scenario is considered. However, our analysis and study can be used for downlink i.e., CR BS to SUs transmission as well. We assume that maximum buffer sizes for DS and BE services are $Q^{(d)}$ and $Q^{(b)}$, respectively.

Let us use $b_n^{(d)}$ and $b_n^{(b)}$ to denote the number of packets in the buffer of DS and BE services, respectively at time slot n.

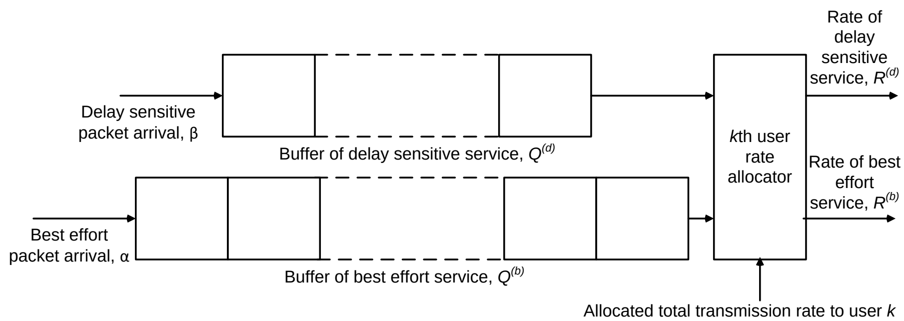

Figure 3.1: Rate allocation for multi-class service transmission for kth SU.

A particular channel is assumed to be time varying which can be modeled using a FSMC as described in Section 2.1.3. Let us use $c_n^{(i,k)} \in \mathcal{C}$ to denote the channel state of ith channel for kth SU at time slot n. Packet transmission rate for ith channel of kth SU at state z is, $\gamma_z^{(i,k)} = rz$ where $0 \leq z \leq (Z-1)$. For each class of service, we consider that random packet arrival process follows batch Bernoulli distribution as described in Section 2.1.5. This batch Bernoulli model for DS packets can be described by $\boldsymbol{\beta} = [\beta_0, \beta_1, ..., \beta_N]$, where N is the maximum number of packets that can arrive for DS service at a time slot. Average packet arrival rate of DS service can be calculated as, $\beta' = \sum_{j=0}^{N} j\beta_j$. Similarly, packet arrival for BE service can be defined by $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, ..., \alpha_M]$, where M is the maximum number of BE packets that can arrive at time slot n. Average packet arrival rate for BE service is, $\alpha' = \sum_{l=0}^{M} l\alpha_l$. We assume that packet(s) arrived during time slot n at a given buffer can be served at time slot $n+1$ at the earliest.

3.1.2 Joint State Considering PUs' Activity, Channel Fading State and Channel Scheduling

We consider a homogeneous system where all SUs have identical packet arrival rates and experience identical but independent channel fading processes. Therefore, we consider a particular SU referred to as tagged user in our study and drop the index for SU, k for notational simplicity. Let us use $m_n^{(i)} \in \{0,1\}$ to denote whether ith channel is assigned to the tagged user or not at time slot n. $m_n^{(i)} = 0$ denotes that ith channel is not assigned to the tagged user at time slot n and vice versa. The joint system state for ith channel at time slot n considering channel fading state ($c_n^{(i)}$), PUs' channel occupancy ($o_n^{(i)}$) and the tagged user channel assignment ($m_n^{(i)}$) can be written as follows: $\mathbf{\Psi}^{(i)} \equiv \{(c_n^{(i)}, o_n^{(i)}, m_n^{(i)}) | 0 \le c_n^{(i)} \le Z-1; 0 \le o_n^{(i)} \le 1; 0 \le m_n^{(i)} \le 1\}$. Let us define a new state variable ξ_n as follows: $\xi_n = \sum_{1 \le i \le N_p} c_n^{(i)} o_n^{(i)} m_n^{(i)}$ and corresponding state space, $\mathbf{\Theta} = \{0, 1, ..., N_p(Z-1)\}$. Using the procedure mentioned in [18], one can obtain transition probability matrix $\mathbf{S}$ for the state space $\mathbf{\Theta}$. From ξ_n one can obtain transmission rate of the tagged SU at a particular time slot n, $R_n = r\xi_n$ where $0 \le R_n \le Y$ and Y is the maximum transmission rate given by $Y = rN_p(Z-1)$.

Now the overall state including buffer state of the system can be denoted as follows: $s^{(u)} \equiv \{(b_n^{(b)} = j, b_n^{(d)} = l, R_n = q) | 0 \le j \le Q^{(b)}; 0 \le l \le Q^{(d)}; 0 \le q \le Y\}$ where u is the state index i.e., $u = j \times (Q^{(d)}+1) \times (Y+1) + l \times (Y+1) + q$ and $0 \le u \le U$, $U = (Q^{(b)}+1) \times (Q^{(d)}+1) \times (Y+1)$. Let us use $\mathcal{S}$ to denote the state space where $\mathcal{S} = \{s^{(0)}, s^{(1)}, ..., s^{(U)}\}$. Since, the state variables can take only discrete value and transition among states can be occurred at the slot boundary, we can view the system as a discrete-time Markov chain (DTMC).

3.2 Problem Description and Rate Allocation Mechanisms

At time slot n using the max rate channel scheduling mechanism, the central scheduler e.g., the scheduler at the CR BS assigns a certain number of empty channels to the tagged SU which can transmit total R_n packet(s). The tagged user's rate allocator needs to allocate this R_n packet(s) between DS and BE services at time slot n. As such, QoS requirements of DS and BE services are maintained. In this section, we study two rate allocation mechanisms. Without loss of generality, let us use $R_n^{(d)}$ and $R_n^{(b)}$ to denote the number of packet(s) to be transmitted from DS and BE service buffers, respectively at time slot n where $R_n^{(d)} + R_n^{(b)} \leq R_n$.

3.2.1 Optimal Rate Allocation Mechanism

In order to find the optimal rate allocation mechanism, the rate allocation problem is formulated as a MDP. As mentioned in Chapter 2, MDP can be described through a set of states $\mathcal{S} = \{s^{(0)}, s^{(1)}, ..., s^{(U)}\}$, a set of actions $\mathcal{A}$, a set of immediate costs/rewards that depend on states and actions and a set of state and action dependent transition probabilities. For the system model presented in this chapter, the set of states is described in the previous section. We assume that at a particular time slot n, the set of all available actions is, $\mathcal{A} = \{a^{(0)}, a^{(r)}, a^{(2r)}, ..., a^{(Y)}\}$ where $a^{(t)}$ represents that the rate allocator of the tagged SU allocates t packet(s) to be transmitted from the DS service buffer i.e., $R_n^{(d)} = t$. The remaining packet(s) is allocated to be transmitted from the BE service buffer i.e., $R_n^{(b)} = R_n - R_n^{(d)}$. t is a feasible action if $Q^{(d)} \geq t$ and $R_n \geq t$. For a given action, transition probabilities between states depend on buffers' dynamics as well as transmission rate, R_n. For an action a_n, at time slot n, the transition probability from state $s^{(j)}$ to state $s^{(l)}$ can be denoted as follows: $P(s_{n+1} = s^{(l)} | s_n = s^{(j)}, a_n = a^{(t)})$ where $a^{(t)} \in \mathcal{A}_{s^{(j)}}$ and $\mathcal{A}_{s^{(j)}}$ is the set of all feasible actions at state $s^{(j)}$. The detailed procedure for calculating transition probabilities is mentioned in Appendix A. Let us use $x(s^{(u)}, a^{(t)})$ to denote the steady state probability of taking action $a^{(t)}$ in state $s^{(u)}$ where $s^{(u)} \in \mathcal{S}$ and $a^{(t)} \in \mathcal{A}_{s^{(u)}}$. Let us use $d^{(d,o)}(s^{(u)}, a^{(t)})$ to denote the queuing delay of DS packets at state $s^{(u)}$ when action $a^{(t)}$ is taken. According to the well known Little's law [36], the average queuing delay, $d^{(d,o)}(s^{(u)}, a^{(t)})$ can be expressed as follows

$$d^{(d,o)}(s^{(u)}, a^{(t)}) = \frac{\left\lfloor \frac{u \mod \{(Q^{(d)}+1)(Y+1)\}}{Y+2} \right\rfloor}{\beta'}, \tag{3.1}$$

where $\lfloor x \rfloor$ denotes floor of x. Using eq. (3.1), one can obtain the queuing delay of DS packets for all states and corresponding feasible actions. For notational convenience let us denote those delays in a vector form as follows

$$\mathbf{d}^{(d,o)}(\mathcal{S},\mathcal{A}) = \begin{bmatrix} d^{(d,o)}(s^{(0)},a^{(0)}) \\ d^{(d,o)}(s^{(1)},a^{(0)}) \\ \vdots \\ d^{(d,o)}(s^{(U)},a^{(0)}) \\ d^{(d,o)}(s^{(0)},a^{(r)}) \\ d^{(d,o)}(s^{(1)},a^{(r)}) \\ \vdots \\ d^{(d,o)}(s^{(U)},a^{(r)}) \\ \vdots \\ d^{(d,o)}(s^{(0)},a^{(Y)}) \\ d^{(d,o)}(s^{(1)},a^{(Y)}) \\ \vdots \\ d^{(d,o)}(s^{(U)},a^{(Y)}) \end{bmatrix}. \tag{3.2}$$

Packet loss due to the buffer over flow occurs if a packet finds the given buffer of the tagged user is full at the time of its arrival. Considering the buffer states that can lead to buffer overflow and associated packet arrival probabilities, one can obtain packet loss probability. Let us use $p_{\text{loss}}^{(b,o)}(s^{(u)},a^{(t)})$ and $p_{\text{loss}}^{(d,o)}(s^{(u)},a^{(t)})$ to denote the packet loss probability of BE and DS services, respectively due to buffer overflow at state $s^{(u)}$ when action $a^{(t)}$ is taken. The packet loss probability of DS service, $p_{\text{loss}}^{(d,o)}(s^{(u)},a^{(t)})$ can be expressed as follows

$$p_{\text{loss}}^{(d,o)}(s^{(u)}, a^{(t)}) = \sum_{i=Q^{(d)}-(\lfloor \frac{u \mod \{(Q^{(d)}+1)(Y+1)\}}{Y+2} \rfloor - a^{(t)})+1}^{M} \beta_i. \tag{3.3}$$

The packet loss probability of DS service for all states and corresponding feasible actions can be written in a vector form as follows

$$\mathbf{p}_{\text{loss}}^{(d,o)}(\mathcal{S}, \mathcal{A}) = \begin{bmatrix} p_{\text{loss}}^{(d,o)}(s^{(0)}, a^{(0)}) \\ p_{\text{loss}}^{(d,o)}(s^{(1)}, a^{(0)}) \\ \vdots \\ p_{\text{loss}}^{(d,o)}(s^{(U)}, a^{(0)}) \\ p_{\text{loss}}^{(d,o)}(s^{(0)}, a^{(r)}) \\ p_{\text{loss}}^{(d,o)}(s^{(1)}, a^{(r)}) \\ \vdots \\ p_{\text{loss}}^{(d,o)}(s^{(U)}, a^{(r)}) \\ \vdots \\ p_{\text{loss}}^{(d,o)}(s^{(0)}, a^{(Y)}) \\ p_{\text{loss}}^{(d,o)}(s^{(1)}, a^{(Y)}) \\ \vdots \\ p_{\text{loss}}^{(d,o)}(s^{(U)}, a^{(Y)}) \end{bmatrix}. \tag{3.4}$$

Similarly, the packet loss probability of BE service, $p_{\text{loss}}^{(b,o)}(s^{(u)}, a^{(t)})$ can be expressed in a vector form as follows

$$p_{\text{loss}}^{(b,o)}(s^{(u)}, a^{(t)}) = \sum_{j=Q^{(b)}-(\lfloor \frac{u}{\{(Q^{(d)}+1)(Y+1)\}} \rfloor - \min\{\lfloor \frac{u}{\{(Q^{(d)}+1)(Y+1)\}} \rfloor, Y-a^{(t)}\})+1}^{N} \alpha_j. \tag{3.5}$$

The packet loss probability of BE service for all states and corresponding feasible actions can be expressed as follows

$$\mathbf{p}_{\text{loss}}^{(b,o)}(\mathcal{S}, \mathcal{A}) = \begin{bmatrix} p_{\text{loss}}^{(b,o)}(s^{(0)}, a^{(0)}) \\ p_{\text{loss}}^{(b,o)}(s^{(1)}, a^{(0)}) \\ \vdots \\ p_{\text{loss}}^{(b,o)}(s^{(U)}, a^{(0)}) \\ p_{\text{loss}}^{(b,o)}(s^{(0)}, a^{(r)}) \\ p_{\text{loss}}^{(b,o)}(s^{(1)}, a^{(r)}) \\ \vdots \\ p_{\text{loss}}^{(b,o)}(s^{(U)}, a^{(r)}) \\ \vdots \\ p_{\text{loss}}^{(b,o)}(s^{(0)}, a^{(Y)}) \\ p_{\text{loss}}^{(b,o)}(s^{(1)}, a^{(Y)}) \\ \vdots \\ p_{\text{loss}}^{(b,o)}(s^{(U)}, a^{(Y)}) \end{bmatrix}. \tag{3.6}$$

The steady state probability for all states and corresponding feasible actions can be expressed in a vector form as follows

$$\mathbf{x}(\mathcal{S}, \mathcal{A}) = \begin{bmatrix} x(s^{(0)}, a^{(0)}) \\ x(s^{(1)}, a^{(0)}) \\ \vdots \\ x(s^{(U)}, a^{(0)}) \\ x(s^{(0)}, a^{(r)}) \\ x(s^{(1)}, a^{(r)}) \\ \vdots \\ x(s^{(U)}, a^{(r)}) \\ \vdots \\ x(s^{(0)}, a^{(Y)}) \\ x(s^{(1)}, a^{(Y)}) \\ \vdots \\ x(s^{(U)}, a^{(Y)}) \end{bmatrix}. \tag{3.7}$$

The queuing delay of DS packets and packet loss probabilities for both classes of services are considered as costs in our optimal problem formulation. The problem is formulated as a constrained MDP where one type of cost is minimized while keeping other costs below target threshold values [57]. In particular, the objective of the optimal rate allocation mechanism is to minimize the average queuing delay of DS packets while keeping the packet loss probabilities of BE as well as DS services below their target thresholds. As such, we can formulate the optimal rate allocation mechanism as follows

[62], [55]:

$$\underset{\mathbf{x}(\mathcal{S},\mathcal{A})}{\text{minimize}}[\mathbf{d}^{(d,o)}(\mathcal{S},\mathcal{A})]^{\intercal}\mathbf{x}(\mathcal{S},\mathcal{A}) \tag{3.8}$$

subject to:

$$[\mathbf{p}_{\text{loss}}^{(b,o)}(\mathcal{S},\mathcal{A})]^{\intercal}\mathbf{x}(\mathcal{S},\mathcal{A}) \leq p_{th}^{(b)} \tag{3.9}$$

$$[\mathbf{p}_{\text{loss}}^{(d,o)}(\mathcal{S},\mathcal{A})]^{\intercal}\mathbf{x}(\mathcal{S},\mathcal{A}) \leq p_{th}^{(d)} \tag{3.10}$$

$$\sum_{a^{(t)}\in\mathcal{A}_{s^{(k)}}} x(s^{(k)},a^{(t)}) = \sum_{s^{(u)}\in\mathcal{S},a^{(t)}\in\mathcal{A}_{s^{(u)}}} x(s^{(u)},a^{(t)})P(s^{(k)}|s^{(u)},a^{(t)}), \quad \forall s^{(k)}\in\mathcal{S} \tag{3.11}$$

$$\sum_{s^{(u)}\in\mathcal{S},a^{(t)}\in\mathcal{A}_{s^{(u)}}} x(s^{(u)},a^{(t)}) = 1 \tag{3.12}$$

$$x(s^{(u)},a^{(t)}) \geq 0, \forall s^{(u)}\in\mathcal{S}, \forall a^{(t)}\in\mathcal{A}_{s^{(u)}}. \tag{3.13}$$

The inequality constraint of eq. (3.9) is to keep the packet loss probability of BE service below a threshold value, $p_{th}^{(b)}$. Similarly, the inequality constraint of eq. (3.10) is to keep the packet loss probability of DS service below a threshold value, $p_{th}^{(d)}$. The equality constraint of eq. (3.11) represents the well known Chapman-Kolmogorov equation whereas eq. (3.12) ensures that sum of all probabilities, $x(s^{(u)}, a^{(t)})$ is equal to one. $[.]^\intercal$ represents transpose of the vector. One can solve the optimization problem in eqs. (3.8)-(3.13) using the well known LP technique [62], [57], [63] and can obtain optimal rate allocation policies, $\mathbf{x}^*(\mathcal{S}, \mathcal{A})$. $x^*(s^{(u)}, a^{(t)})$ denotes the steady state probability of taking action $a^{(t)}$ in state $s^{(u)}$ that minimizes the average queuing delay of DS packets while satisfies packet loss probability constraints.

We used MATLAB [63] to solve the LP formulated in eqs. (3.8)-(3.13). After obtaining the optimal values of $\mathbf{x}^*(\mathcal{S}, \mathcal{A})$, one can measure average queuing delay and packet loss probabilities with the optimal rate allocation mechanism as follows.

Packet Loss Probability

Packet loss probability corresponding to the optimal action $\mathbf{x}^*(\mathcal{S}, \mathcal{A})$ can be obtained as follows

$$P_{\text{loss}}^{(b,o)^*} = [\mathbf{p}_{\text{loss}}^{(b,o)}(\mathcal{S}, \mathcal{A})]^\intercal \mathbf{x}^*(\mathcal{S}, \mathcal{A}), \tag{3.14}$$

$$P_{\text{loss}}^{(d,o)^*} = [\mathbf{p}_{\text{loss}}^{(d,o)}(\mathcal{S}, \mathcal{A})]^\intercal \mathbf{x}^*(\mathcal{S}, \mathcal{A}), \tag{3.15}$$

where $P_{\text{loss}}^{(b,o)^*}$ represents the packet loss portability of BE service and $P_{\text{loss}}^{(d,o)^*}$ represents the packet loss portability of DS service corresponding to optimal action $x^*(\mathcal{S}, \mathcal{A})$.

Average Queuing Delay

The average queuing delay of DS packets, $D_{\text{avg}}^{(d,o)^*}$ corresponding to the optimal action $x^*(\mathcal{S},\mathcal{A})$ can be written as follows

$$D_{\text{avg}}^{(d,o)^*} = \frac{[\mathbf{d}^{(d,o)}(\mathcal{S},\mathcal{A})]^\intercal \mathbf{x}^*(\mathcal{S},\mathcal{A})}{1 - P_{\text{loss}}^{(d,o)^*}}. \tag{3.16}$$

3.2.2 Suboptimal Rate Allocation Mechanism

It is well known that the optimal policies for constrained MDP are random [55], [57]. Therefore, implementation of optimal rate allocation policies can be difficult. In what follows, we propose a low-complexity suboptimal rate allocation mechanism and study its performance. According to the suboptimal rate allocation mechanism, the available transmission rate at time slot n of the tagged SU, R_n is allocated between two classes of services as follows.

– If $R_n \leq b_n^{(d)}$, the rate allocator of the tagged user allocates R_n packet(s) to be transmitted from DS service buffer i.e., $R_n^{(d)} = R_n$. In this case, all R_n packet(s) is transmitted from DS buffer and no packet is transmitted from BE buffer.

– On the other hand, if $R_n > b_n^{(d)}$, the rate allocator allocates $b_n^{(d)}$ packet(s) to be transmitted from DS service buffer i.e., $R_n^{(d)} = b_n^{(d)}$. The remaining packet(s) is allocated to the BE service i.e., $R_n^{(b)} = R_n - R_n^{(d)}$.

In order to analyze the performance of suboptimal rate allocation mechanism, we develop a queuing analytic model. In the developed queuing analytic model, the state transition probabilities for infinite buffer space for both buffers can be expressed using eq. (3.17). Now we can model eq. (3.17) as a QBD process by making a block of sub matrices. For the finite buffer case with level $X = \lfloor \frac{Q^{(b)}}{Y} \rfloor$, eq. (3.17) can be expressed as follows

$$\mathbf{P}^{(s)} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ X-1 \\ X \end{array} \begin{bmatrix} \mathbf{A} & \mathbf{C} & & & & & \\ \mathbf{D} & \mathbf{E} & \mathbf{F_0} & & & & \\ & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & & \\ & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F'_0} \\ & & & & & \mathbf{F'_2} & \mathbf{F'_1} \end{bmatrix}. \tag{3.18}$$

The definition of inner sub matrices of eq. (3.18) is mentioned in the Appendix B. One can obtain the steady state probability vector $\boldsymbol{\pi}$ of eq. (3.18) using the procedure mentioned in Appendix IV. Let us use π_k to denote the steady state probability corresponding to state k where $k = 0, 1, ..., U$. Using the steady state probabilities, one can measure different packet-level performance parameters for suboptimal mechanism as follows.

Packet Loss Probability

A packet will be lost due to the buffer over flow if it finds the given buffer full upon its arrival. Packet loss probability can be measured from the steady state probabilities of those states that lead to buffer overflow upon arrival

$$
\mathbf{P}_{\text{inf}}^{(\mathbf{s})} = \left[\begin{array}{c|ccccc|ccccc|ccc}
\mathbf{B}_0^{(0)} & \mathbf{B}_{1+}^{(0)} & \cdots & \mathbf{B}_{M+}^{(0)} & & & & & & & & & & \\
\hline
\mathbf{B}_{1-}^{(1)} & \mathbf{B}_0^{(1)} & \mathbf{B}_{1+}^{(1)} & \cdots & \mathbf{B}_{M+}^{(1)} & & & & & & & & & \\
\vdots & \vdots & & & \ddots & & & & & & & & & \\
\mathbf{B}_{(Y-M+1)-}^{(Y-M+1)} & \mathbf{B}_{(Y-M)-}^{(Y-M+1)} & \cdots & \mathbf{B}_0 & \mathbf{B}_{(1)+}^{(Y-M+1)} & \mathbf{B}_{(M-1)+}^{(Y-M+1)} & \mathbf{B}_{M+} & & & & & & & \\
\vdots & \vdots & & & & \vdots & \vdots & \ddots & & & & & & \\
\mathbf{B}_{Y-}^{(Y)} & \mathbf{B}_{(Y-1)-}^{(Y)} & \cdots & \cdots & \mathbf{B}_{(1)-}^{(Y)} & \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & & & & \\
\hline
 & \mathbf{B}_{Y-} & \cdots & \cdots & \cdots & \mathbf{B}_{1-} & \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & & & \\
 & & \ddots & & & \vdots & \vdots & & & \ddots & & & & \\
 & & & \mathbf{B}_{Y-} & & \mathbf{B}_{(Y-M+1)-} & \mathbf{B}_{(Y-M)-} & \cdots & \cdots & \cdots & \mathbf{B}_{(M-1)+} & \mathbf{B}_{M+} & & \\
 & & & & \ddots & \vdots & \vdots & & & \ddots & \vdots & \vdots & \ddots & \\
 & & & & & \mathbf{B}_{Y-} & \mathbf{B}_{(Y-1)-} & \cdots & \cdots & \mathbf{B}_{1-} & \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} \\
\hline
 & & & & \ddots & & & & & \ddots & & & & \ddots
\end{array}\right], \tag{3.17}
$$

of packet(s) and corresponding packet arrival probabilities. The packet loss probability due to buffer overflow of BE service can be expressed as

$$P_{\text{loss}}^{(b,s)} = \sum_{b^{(b)}=(Q^{(b)}-M+1)}^{Q^{(b)}} p(b^{(b)}) \times \sum_{i=(Q^{(b)}-b^{(b)}+1)}^{M} \alpha_i, \tag{3.19}$$

where $p(b^{(b)})$ corresponds to the steady state probability of having $b^{(b)}$ packet(s) in the buffer of BE service of the tagged user and is obtained as follows

$$p(b^{(b)}) = \boldsymbol{\pi}(b^{(b)}, b^{(d)}, R_n)\mathbf{1}, \qquad 0 \leq b^{(b)} \leq Q^{(b)}, \tag{3.20}$$

where $\mathbf{1}$ is a column vector of length $(Q^{(d)} + 1)Z$. Similarly, the probability that the DS buffer has $b^{(d)}$ packet(s) can be written as

$$p(b^{(d)}) = \boldsymbol{\pi}(b^{(b)}, b^{(d)}, R_n)\mathbf{1}, \qquad 0 \leq b^{(d)} \leq Q^{(d)}, \tag{3.21}$$

where $\mathbf{1}$ is a column vector of length $(Q^{(b)} + 1)Z$. Using eq. (3.21), the packet loss probability due to buffer overflow of DS service can be written as

$$P_{\text{loss}}^{(d,s)} = \sum_{b^{(d)}=(Q^{(d)}-N+1)}^{Q^{(d)}} p(b^{(d)}) \times \sum_{i=(Q^{(d)}-b^{(d)}+1)}^{N} \beta_i. \tag{3.22}$$

Average Queuing Delay

The average queuing delay of DS packets can be written as

$$D_{\text{avg}}^{(d,s)} = \frac{\sum_{b^{(d)}=1}^{Q^{(d)}} p(b^{(d)}) \times b^{(d)}}{\beta' \times (1 - P_{\text{loss}}^{(d,s)})}. \tag{3.23}$$

Delay Distribution

As mentioned, DS services e.g., voice and video require the data streams to be delivered within a certain time limit [64]. For example, a DS packet requires to be delivered within $D_{\text{max}}^{(d,s)}$ time slots in order to maintain interactive and streaming nature. From the cumulative distribution function (CDF) of delay, one can obtain the probability of transmitting the packet within $D_{\text{max}}^{(d,s)}$ time slots i.e., $F_D^{(d,s)}(D_{\text{max}}^{(d,s)}) = \Pr[D \leq D_{\text{max}}^{(d,s)}]$ where D is the packet delay in time slots. Also, a DS packet can have a statistical delay guarantee bound i.e., $\Pr[D > D_{\text{max}}^{(d,s)}] \leq \epsilon$ where ϵ represents the QoS violation upper bound [65]. The statistical delay guarantee can be expressed in terms of CDF of delay as follows

$$F_D^{(d,s)}(D_{\text{max}}^{(d,s)}) \geq 1 - \epsilon. \tag{3.24}$$

Therefore, delay distribution is one of the most important QoS parameters for DS packets. In fact, statistical delay guarantee of packets for DS service is more relevant for CRN as it is difficult to maintain if not impossible, hard delay guarantee. In this section, we describe the procedure for calculating CDF of delay. The delay distribution/CDF of a DS packet can be derived from the so-called absorbing Markov chain [18]. According to the absorbing Markov chain, the Markov chain starts from the state that corresponds to the arrival of a DS packet at the DS buffer. The packet gets absorbed when it reaches at the front of the buffer where it will get the next turn for transmission. In this case, there is no arrival in the DS buffer of the tagged user while the process moves towards the absorbing state i.e., zero packet at the DS buffer. The absorbing Markov chain, P_{abs} can be obtained from eq. (3.18) by setting $\beta_0 = 1$ and $\beta_i = 0$. The delay distribution can be determined from the initial probability vector i.e., the probability vector associated with the number packets including arriving packet and the number of packets ahead of the arriving packet in the DS buffer. This initial probability vector, $\boldsymbol{\zeta}^{(0)}$ can be obtained as follows

$$\boldsymbol{\zeta}^{(0)} = [\boldsymbol{\zeta}_0^{(0)} \quad \boldsymbol{\zeta}_1^{(0)} \quad \cdots \boldsymbol{\zeta}_{Q^{(b)}}^{(0)}], \tag{3.25}$$

where $\boldsymbol{\zeta}_i^{(0)} = [\boldsymbol{\zeta}_{i,0}^{(0)} \quad \boldsymbol{\zeta}_{i,1}^{(0)} \quad \cdots \quad \boldsymbol{\zeta}_{i,Q^{(d)}}^{(0)}]$ and $\boldsymbol{\zeta}_i^{(0)}$ can be be derived as follows

$$\boldsymbol{\zeta}_{i,0}^{(0)} = \mathbf{0}, \tag{3.26}$$

where $\mathbf{0}$ is a row vector with length $Y+1$.

$$\boldsymbol{\zeta}_{i,l}^{(0)} = \begin{cases} \frac{1}{1-\beta_0}[\sum_{j=1}^{l}\sum_{k=1}^{j}\sum_{0\le h\le N_pZ}\frac{\beta_j}{j}\{\boldsymbol{\pi}_{i,l-k+rh}\mathbf{U}^{(h)} \\ +\boldsymbol{\pi}_{i,l-k+rh}\sum_{\lceil\frac{h+1}{r}\rceil\le m\le N_pZ}\mathbf{U}^{(m)}\}], \quad \text{if } 1\le l\le N \\ \frac{1}{1-\beta_0}[\sum_{j=1}^{N}\sum_{k=1}^{j}\sum_{0\le h\le N_pZ}\frac{\beta_j}{j}\boldsymbol{\pi}_{i,l-k+rh}\mathbf{U}^{(h)}], \\ \quad \text{if } N\le l\le Q^{(d)}-Y+1 \\ \frac{1}{1-\beta_0}[\sum_{j=1}^{N}\sum_{k=1}^{j}\sum_{0\le h\le Q^{(d)}-l+k}\frac{\beta_j}{j}\boldsymbol{\pi}_{i,l-k+h} \\ \mathbf{U}^{(\frac{h}{r})}], \quad \text{if } Q^{(d)}-Y+1\le l\le Q^{(d)}-1 \\ \sum_{l=Q^{(d)}}^{Q^{(d)}+N}\frac{1}{1-\beta_0}[\sum_{j=l-Q^{(d)}+1}^{N} \\ \sum_{k=l-Q^{(d)}+1}^{j}\sum_{h=0}^{Q^{(d)}-l+k}\frac{\beta_j}{j}\boldsymbol{\pi}_{i,l-k+h}\mathbf{U}^{(\frac{h}{r})}], \quad \text{otherwise} \end{cases} \tag{3.27}$$

where $\boldsymbol{\pi}_{i,l} = [\pi(b^{(b)}=i, b^{(d)}=l, R=0) \quad \pi(b^{(b)}=i, b^{(d)}=l, R=1) \quad \cdots \pi(b^{(b)}=i, b^{(d)}=l, R=Y)]$, $0\le i\le Q^{(b)}$, $0\le l\le Q^{(d)}$.

We consider that a DS packet will be transmitted within D (time slots) where $0\le D\le\infty$. The CDF of delay for DS packet can be written as follows

$$F_D^{(d,s)} = \sum_{i=0}^{Q^{(b)}}\boldsymbol{\zeta}_{i,0}^{(D)}\mathbf{1}, \tag{3.28}$$

where $\mathbf{1}$ is a column vector with length $Y+1$ and $\boldsymbol{\zeta}^{(D)} = \boldsymbol{\zeta}^{(0)}P_{abs}^{D}$ where P_{abs}^{D} is the transition probability matrix derived from P_{abs} by multiplying P_{abs}

with itself for D times.

3.3 Numerical Results

In this section, we have presented selected numerical results and compared the performances of optimal and suboptimal rate allocation mechanisms. Presented numerical results are validated via computer simulation using MATLAB. We have also compared the performances of these rate allocation mechanisms with the well known first-in and first-out (FIFO) packet scheduling mechanism. According to this FIFO mechanism, packets are served based on their arrival times. For example, if a BE packet arrives before a DS packet, the BE packet will be transmitted before the DS packet. Presented results of FIFO packet scheduling mechanism and the delay distribution of optimal rate allocation mechanism are obtained via simulation. In the numerical examples, we assume that the number of channel states, $Z = 3$, buffer size of DS service, $Q^{(d)} = 5$, buffer size of BE service, $Q^{(b)} = 27$, average packet arrival rate of BE service, $\alpha' = 0.35$, the number of channels, $N_p = 2$, Nakagami fading parameter, $m = 1.1$ and $r = 1$.

In Fig. 3.2, we have plotted average queuing delay of DS packets, $D_{\text{avg}}^{(d,o)^*}$ versus target packet loss probability of BE packets, $p_{th}^{(b)}$ using the optimal rate allocation mechanism for different number of SUs, K in the system. In this case, we consider that the average packet arrival rate of DS service, $\beta' = 0.3$ and PUs' activity, $\rho = 0.5$. For a particular number of SUs, if the target

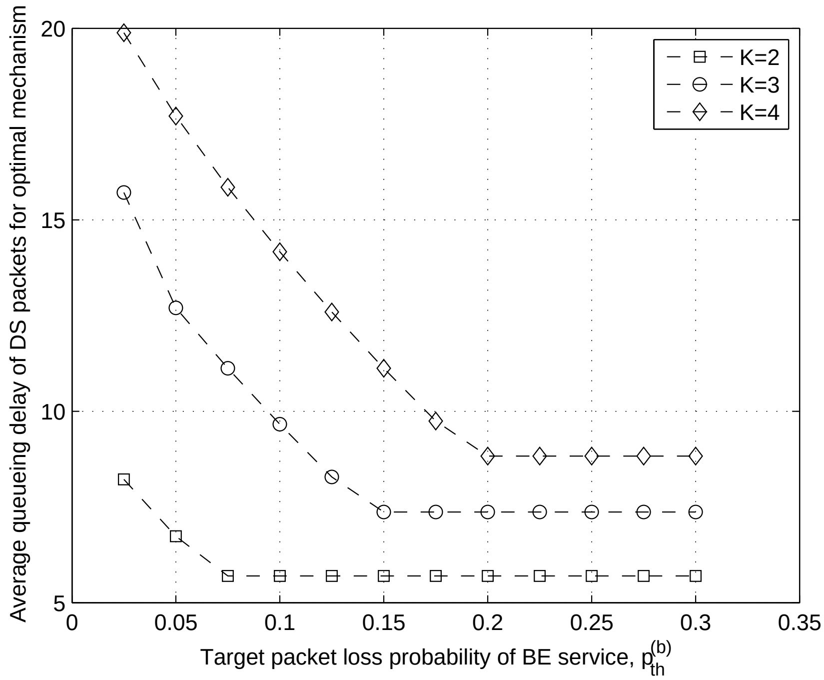

Figure 3.2: Average queuing delay of DS packets vs. target packet loss probability of BE service.

packet loss probability of BE service, $p_{th}^{(b)}$ increases, the average queuing delay of DS packets decreases. The reason can be explained as follows. When the target packet loss probability of BE service, $p_{th}^{(b)}$ is lower, in order to satisfy the lower $p_{th}^{(b)}$, kth SU's rate allocator overall allocates more packets to be transmitted from the BE service buffer out of the available transmission rate, R_n. Therefore, the queuing delay of DS packets is higher. On the other hand, for a higher target packet loss probability of BE service, $p_{th}^{(b)}$, kth SU's rate allocator overall allocates less packets to be transmitted from the BE service buffer which leads to a lower average delay of DS packets. However, the average queuing delay of DS packets does not decrease after a particular value even if the target packet loss probability of BE service is further increased. This particular value is the lowest possible value of average queuing delay of DS packets for given operating parameters and the number of SUs in the system. Since the main target of the optimal rate allocation mechanism is to minimize queuing delay of DS packets, for the rest of numerical results, we set the lowest possible value of $p_{th}^{(b)}$ as the target packet loss probability of BE service in eq. (3.9) that corresponds to the minimum value of $D_{\text{avg}}^{(d,o)^*}$ for given number of SUs in the system.

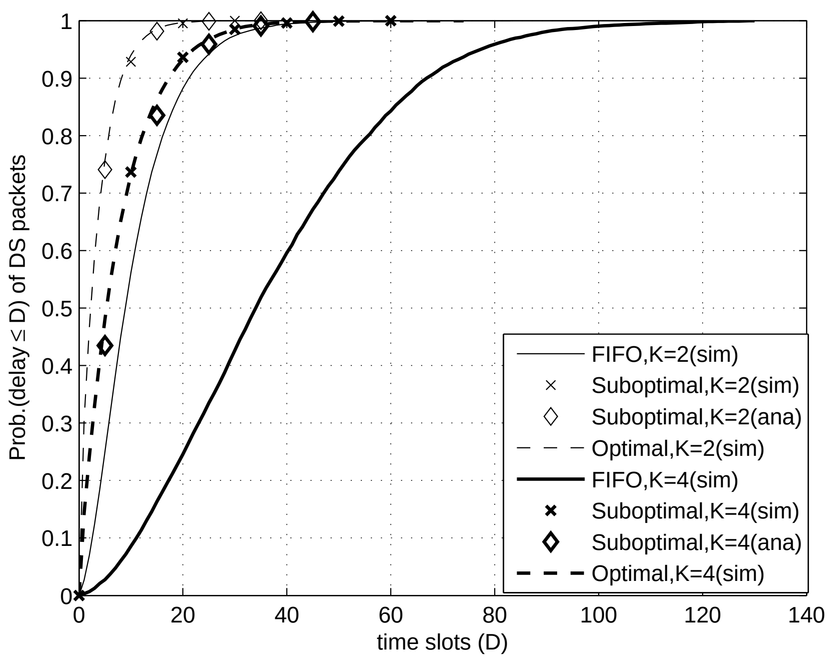

Figure 3.3: Effect of number of SUs (K) on the delay distribution of DS packets for different rate allocation mechanisms (sim=simulation, ana=analysis).

In Fig. 3.3, we have plotted the effect of number of SUs, K on the delay distribution of DS packets for different rate allocation mechanisms where we consider that the average packet arrival rate of DS service, $\beta' = 0.3$ and PUs' activity, $\rho = 0.5$. As expected, FIFO packet scheduling mechanism offers longer queuing delay for DS packets compared to other schemes. This can be explained as follows. With the FIFO packet scheduling mechanism, DS packets are not given priority rather they are transmitted based on their times of arrival. Therefore, a DS packet experiences a longer delay compared to other schemes. Fig. 3.3 also shows that the suboptimal rate allocation mechanism offers a packet delay distribution quite similar to the optimal rate allocation mechanism for given operating parameters as well as the number of SUs in the system. The reason is as follows. According to the suboptimal mechanism, the rate allocator always gives higher priority to DS service and serves DS packets as soon as possible based on the available transmission rate. Therefore, the suboptimal rate allocation mechanism also achieves the lowest possible minimum delay for DS packets.

Since the number of available channels is limited, the transmission opportunity of the tagged SU reduces when the number of SUs in the system increases. Therefore, the delay performance for all schemes degrades for a higher number of SUs as shown in Fig. 3.3. Similarly, higher PUs' activity, ρ reduces transmission opportunities of SUs. Therefore, higher PUs' activity, ρ deteriorates the delay performance for all schemes as shown in Fig. 3.4. In this case, we consider that the average packet arrival rate of DS service, $\beta' = 0.3$ and number of SUs, $K = 2$. Fig. 3.5 shows the delay performance for different packet arrival rate of DS service, β'. As expected, higher

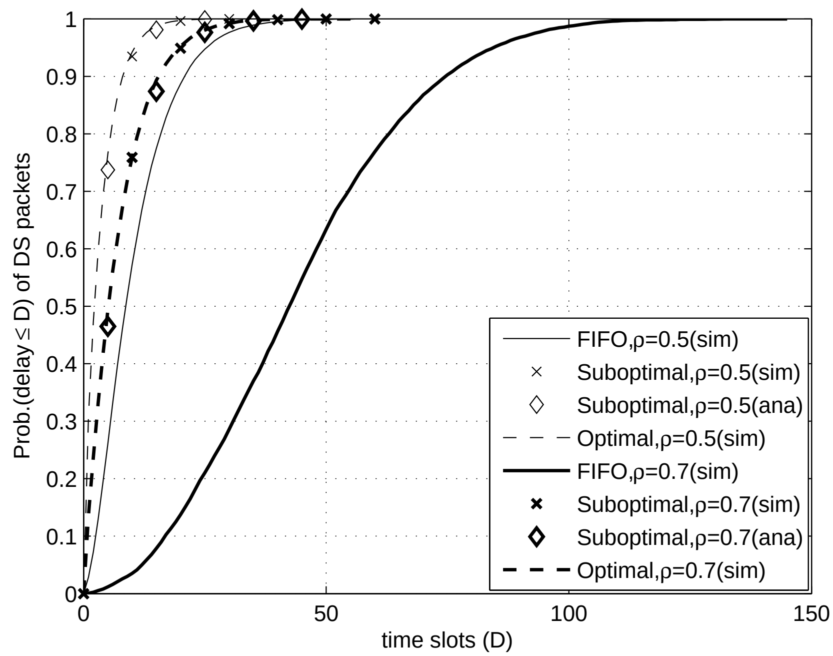

Figure 3.4: Effect of PUs' activity (ρ) on the delay distribution of DS packets for different rate allocation mechanisms (sim=simulation, ana=analysis).

packet arrival rate of DS service, β' deteriorates the delay performance for all schemes. In this case, we consider that the PUs' activity, $\rho = 0.5$ and number of SUs, $K = 2$.

In Fig. 3.6, we have shown the effect of number of SUs, K and average packet arrival rate of DS service, β' on the packet loss probability of DS service for different rate allocation mechanisms. In this figure we consider that the PUs' activity, $\rho = 0.5$. As expected, FIFO packet scheduling mechanism offers a higher packet loss probability for the DS service while suboptimal rate allocation mechanism offers a quite similar to the optimal rate alloca-

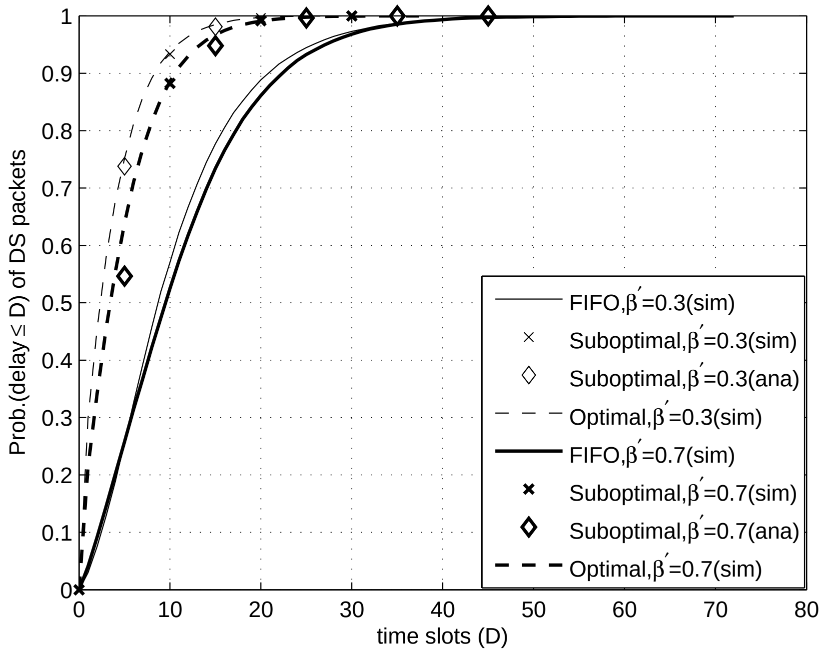

Figure 3.5: Effect of average packet arrival rate of DS service (β') on the delay distribution of DS packets for different rate allocation mechanisms (sim=simulation, ana=analysis).

tion mechanism. Fig. 3.6 also shows that higher number of SUs as well as higher packet arrival rate of DS service increases the packet loss probability of DS service. Similarly, higher PUs' activity, ρ increases the packet loss probability of DS service as shown in Fig. 3.7. In Fig. 3.7, we consider that the average packet arrival rate of DS service, $\beta' = 0.3$.

Fig. 3.8, shows the effect of number of SUs, K and average packet arrival rate of DS service, β' on the packet loss probability of BE service for different rate allocation mechanisms. In this case, FIFO mechanism offers a

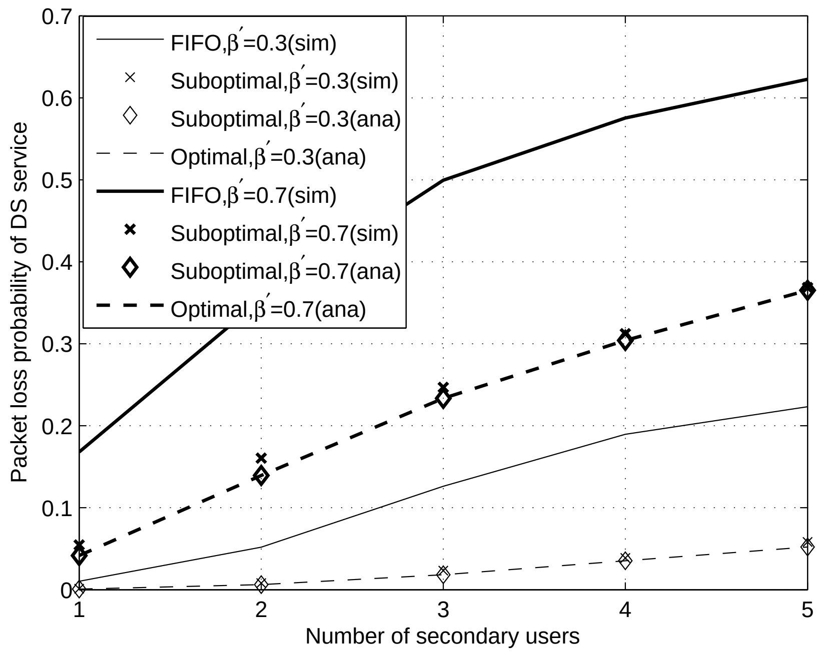

Figure 3.6: Effect of number of SUs, K and average packet arrival rate of DS service (β') on the packet loss probability of DS service.

lower packet loss probability for the BE service compared to other schemes. This can be explained as follows. With the FIFO mechanism, no priority is given for DS service as the packets are served based on their times of arrival. Therefore, the packet loss performance of BE service is improved at the expense of packet loss as well as delay performance of the DS service. Again, the suboptimal rate allocation mechanism has a packet loss probability for BE service quite similar to the optimal rate allocation mechanism. Packet loss probability for all rate allocation mechanisms increases for higher number of SUs in the system as shown in Fig. 3.8. Packet loss probability also increases for higher packet arrival rate of DS service, β' and higher PUs'

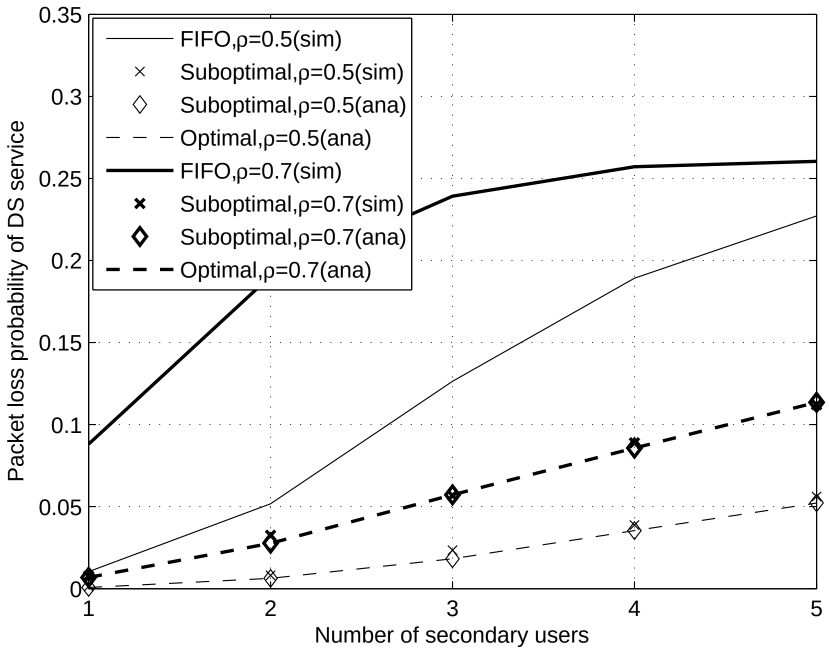

Figure 3.7: Effect of number of SUs (K) and PUs' activity (ρ) on the packet loss probability of DS service (sim=simulation, ana=analysis).

activity, ρ as shown in Figs. 3.8 and 3.9, respectively.

In Fig. 3.10, we have plotted the effect of number of SUs, K and average packet arrival rate of DS service, β' on the average queuing delay of DS packets for different rate allocation mechanisms. In this case, FIFO packet scheduling mechanism offers a longer average queuing delay of DS packets compared to other schemes. On the other hand, the suboptimal rate allocation mechanism has an average queuing delay quite similar to the optimal rate allocation mechanism. The average queuing delay of DS packets increases for higher number of SUs, K, higher packet arrival rate of

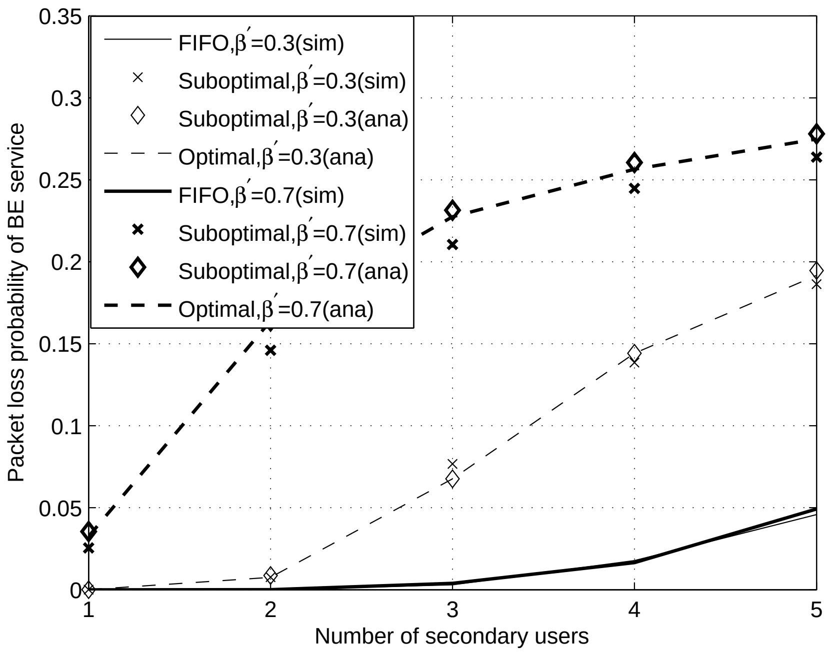

Figure 3.8: Effect of number of SUs (K) and average packet arrival rate of DS service (β') on the packet loss probability of BE service (sim=simulation, ana=analysis).

DS service, β' and higher PUs' activity, ρ as shown in Figs. 3.10 and 3.11, respectively.

All the presented numerical examples show that the suboptimal rate allocation mechanism performs quite similar to the optimal rate allocation mechanism for given operating parameters as well as the number of SUs in the system while it achieves the lowest possible minimum delay for DS packets.

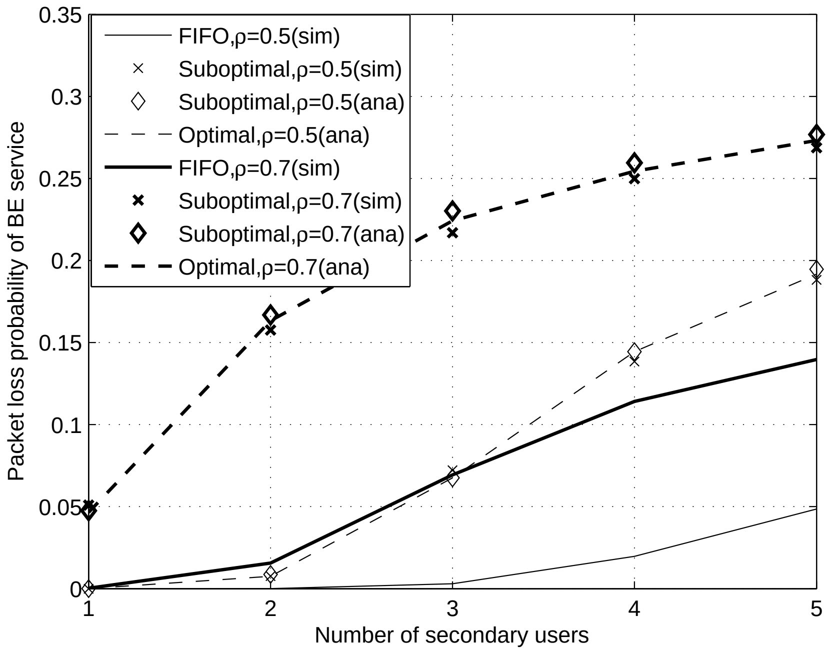

Figure 3.9: Effect of number of SUs (K) and PUs' activity (ρ) on the packet loss probability of BE service (sim=simulation, ana=analysis).

3.4 Application

Since the availability of spectrum is very dynamic in CRN, a rate allocation mechanism may not be sufficient to meet QoS requirements of different classes of services for given operating parameters. In fact, one can observe from Fig. 3.2, in order to satisfy a target packet loss probability of BE service, the CRN can support a certain number of SUs in the system given other operating parameters. Therefore, a proper admission control mechanism is required to maintain QOS requirements for different classes of services. In this section, we provide an example that demonstrates the significance of our

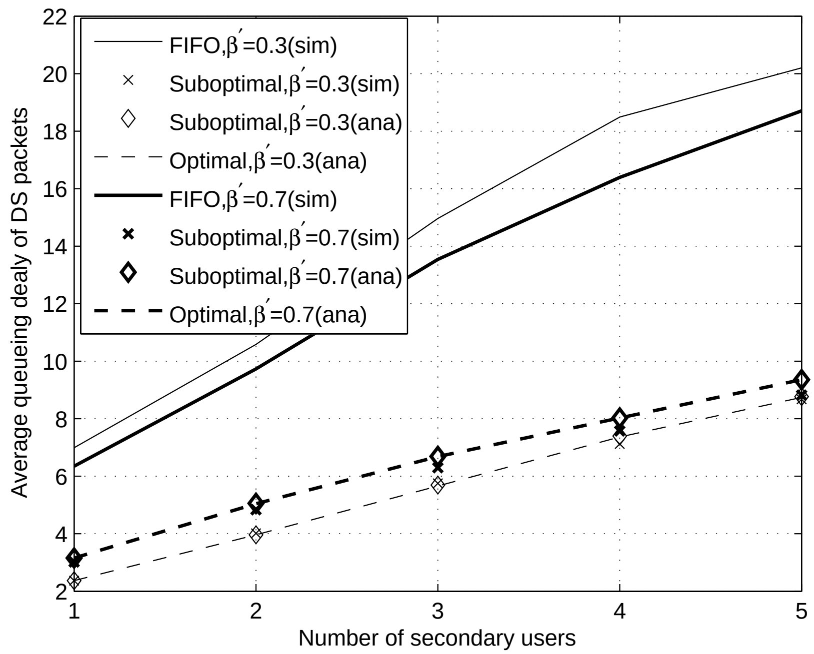

Figure 3.10: Effect of number of SUs (K) and average packet arrival rate of DS service (β') on the average queuing delay of DS packets (sim=simulation, ana=analysis).

Table 3.1: Packet arrival rate β' and primary user activity ρ vs. number of SUs for given QoS requirements ($P_{t,loss}^{(b,s)} \leq 0.05$, $P_{t,loss}^{(d,s)} \leq 0.1$ and $D_{t,max}^{(d,s)} =$ 10 (time slots) with $\epsilon = 0.2$)

β'	ρ	$K_{P_{t,loss}^{(b,s)}}$	$K_{P_{t,loss}^{(d,s)}}$	$K_{D_{t,max}^{(d,s)}}$	K_s
0.3	0.5	2	5	3	2
0.7	0.5	1	1	2	1
0.3	0.7	1	4	1	1

developed queuing analytic model of the suboptimal rate allocation mechanism for admission control decision in CRN.

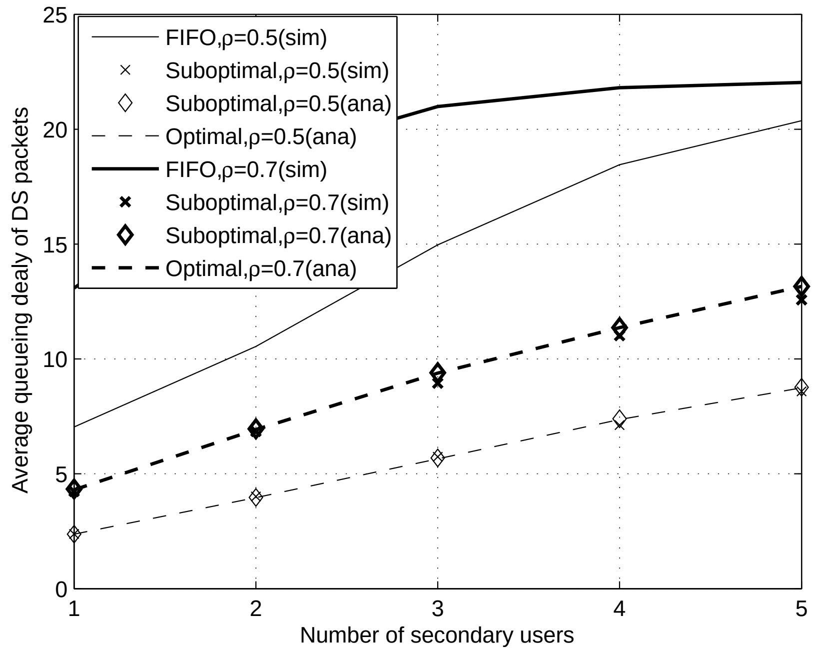

Figure 3.11: Effect of number of SUs (K) and PUs' activity (ρ) on the average queuing delay of DS packets (sim=simulation, ana=analysis).

Let us assume that each SU in the network has a target packet loss probability for DS service, $P_{t,loss}^{(d,s)} \leq 0.1$ and BE service, $P_{t,loss}^{(b,s)} \leq 0.05$. We also consider that each SU has a target statistical delay guarantee for DS packet i.e., $D_{t,max}^{(d,s)} = 10$ (time slots) with probability, $\epsilon = 0.2$. Let us use $K_{D_{t,max}^{(d,s)}}$ to denote the maximum number of SUs that can be supported to maintain target statistical delay guarantee for DS packet. Using our developed analytic model, one can obtain $K_{D_{t,max}^{(d,s)}}$ for different values of average packet arrival rate of DS service, β' and PUs' activity, ρ. Let us use $K_{P_{t,loss}^{(b,s)}}$ and $K_{P_{t,loss}^{(d,s)}}$ to denote the maximum number of SUs that can be supported to maintain $P_{t,loss}^{(b,s)}$ and $P_{t,loss}^{(d,s)}$, respectively. Again, one can obtain $K_{P_{t,loss}^{(b,s)}}$ and $K_{P_{t,loss}^{(d,s)}}$ for different values of average packet arrival rate of DS service, β' and PUs' activity, ρ from the developed analytic model (see Figs. 3.8, 3.9, 3.6 and 3.7). We have listed the values of $K_{P^{(d,s)}_{t,loss}}$, $K_{P^{(b,s)}_{t,loss}}$ and $K_{D^{(d,s)}_{t,max}}$ for different values of β' and ρ in Table 4.1. For given values of β' and ρ, the maximum number of SUs that satisfies all three QoS requirements is equal to minimum of $K_{P^{(d,s)}_{t,loss}}$, $K_{P^{(b,s)}_{t,loss}}$ and $K_{D^{(d,s)}_{t,max}}$ i.e., $K_s = \min(K_{P^{(d,s)}_{t,loss}}, K_{P^{(b,s)}_{t,loss}}, K_{D^{(d,s)}_{t,max}})$ which is also listed in column 6 of the Table 4.1.

From Table 4.1 it is obvious that for $\beta' = 0.3$ and $\rho = 0.5$, CRN can support maximum 2 SUs in order to maintain the target QoS requirements. If the CRN is supporting 2 SUs and a new request is made, the CRN has to reject the request.

Chapter 4

Cross-Layer Performance in Presence of Sensing Errors

[3]As mentioned in Chapter 1, spectrum sensing is one of the most challenging tasks in CRN [66]. Several factors make perfect spectrum sensing practically challenging [67]. For example, CRN may not be able to detect PUs' signal due to multi-path fading and time dispersion of wireless channels [67]. Shadowing effect leads to so-called hidden terminal problem. Moreover, CRN requires hardware that can sample at a high rate over a wide range of frequencies [35]. Also, sensing accuracy depends on the sensing duration. A longer sensing duration increases the reliability of sensing. However, a longer sensing duration reduces transmission time of CRN [68]. Therefore, there exists a trade-off between sensing accuracy and transmission efficiency. As mentioned in Chapter 1 that there are two types of sensing errors, namely miss-detection and false alarm. CRN may not be able to detect an active PU in a given channel. Such sensing error is referred to as a miss-detection.

[3]The research work presented in this chapter has been presented as S M Shahrear Tanzil and Md. Jahangir Hossain, "Cross-layer performance analysis for cognitive radio network with a random transmission protocol in presence of sensing errors", in Proc. of the *Int. Conf. on Cognitive Radio Oriented Wireless Networks (CROWNCOM'13)*, Washington DC, USA, Jul. 2013.

On the other hand, CRN may detect a channel being used by a PU where in reality the channel is idle. This type of sensing error is known as a false alarm. Sensing errors do not only degrade the performance of CRN but also lead to collisions of SUs transmission with that of PUs transmission. Therefore, it is important to analyze the packet-level performance of CRN jointly considering the time varying nature of the communication channel, availability of the channel and channel sensing errors. Authors of [18], [69] developed a queuing model to analyze packet-level performances assuming perfect channel sensing in their models. Recently, Wang *et. al* [70] also developed a Markov model for queuing analysis with sensing errors. However, they considered the traditional deterministic transmission protocol[4] in their Markov model.

In this chapter, we study performance of a random transmission protocol in presence of sensing errors for single class of service. According to the random transmission protocol, SUs access the channel with a certain probability based on the sensing outcome. We develop a queuing analytic model for CRN that considers sensing errors as well as a random transmission protocol. We analyze the developed MC as a QBD process. With the help of our analytic model, one can easily measure different packet-level performance parameters for a given number of SUs, probability of sensing errors, transmission probabilities of CRN and other operating parameters. With the developed model, the collision probability to the PUs can also

[4]By deterministic transmission protocol we mean that SUs access the channel when it is sensed as idle and they do not access the channel when the channel is sensed as busy.

be measured. As such our analysis is useful in making a CAC decision in CRN for a given collision probability to the PUs while maintaining target QoS requirements for SUs. For example, the SUs have certain QoS requirements in terms of delay and packet loss probability. On the other hand, the PUs can tolerate collision with a certain probability. Using our model, one can determine the number of SUs that can be supported while maintaining these QoS requirements for given false alarm and miss-detection probabilities. We also compare the performance of the considered random transmission protocol with the performance of the traditional deterministic transmission protocol. Presented numerical results demonstrate the significance of our analytic model in terms of CAC decision. Numerical results also interestingly show that CRN can support a higher number of SUs using the considered random transmission protocol than the deterministic transmission protocol when transmission probabilities of the random transmission protocol are properly chosen.

The rest of the chapter is organized as follows. In Section 4.1 we describe details of our system model. Section 4.2 describes the random transmission protocol and Section 4.3 presents details of queuing analytic model. While in Section 4.4, we analyze different performance parameters, in Section 4.5, we present some numerical results and application of the developed model.

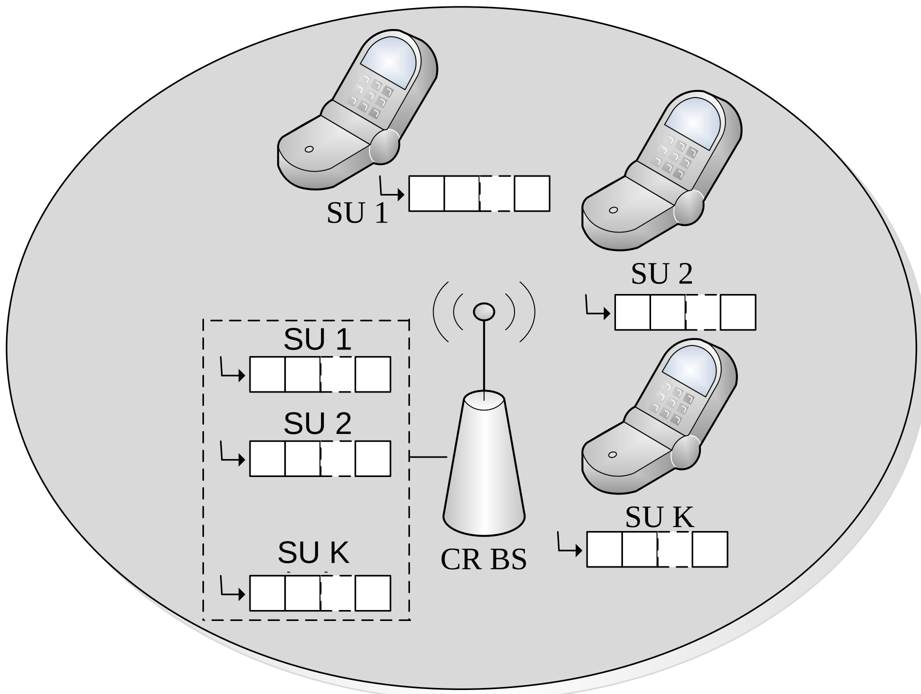

Figure 4.1: An example of infrastructure-based cognitive radio network with data link layer buffers.

4.1 System Model

4.1.1 Overall Description

A single channel infrastructure-based CRN is considered as described in Section 2.1.1. Since the CR BS is supporting K SUs, it is considered that there are K buffers at the CR BS corresponding to K SUs as shown in Fig. 4.1. Our analysis and study can be used for both downlink and uplink transmission. At the beginning of each time slot, CRN senses the channel and assigns the channel to a SU according to the max rate channel scheduling mechanism. Let us use random variable $\hat{O}$ to denote the estimated channel occupancy state where $\hat{O} = 1$ represents the fact that the channel is sensed as idle and $\hat{O} = 0$ represents the fact that the channel is sensed as busy/occupied. Due to the sensing errors as mentioned in Chapter 1, in a given time slot n, the value of $\hat{O}_n$ may not be same as the value of O_n where O_n is the actual channel occupancy at time slot n. In particular, when there is a miss-detection, $\hat{O}_n = 1$ given $O_n = 0$. On the other hand, due to the false alarm, $\hat{O}_n = 0$ given $O_n = 1$.

As described in Chapter 2, the time varying nature of the wireless channel can be modeled as a FSMC. $c_n \in \mathcal{C}$ denotes the channel state at time slot n. We consider a single class of traffic transmission where the packet arrival process at the data link layer follows batch Bernoulli random process described by $\boldsymbol{\beta} = [\beta_0, \beta_1, ..., \beta_N]$, and N is the maximum number of packets that can arrive at a time slot. Average packet arrival rate can be calculated as, $\beta' = \sum_{j=0}^{N} j\beta_j$. We assume that packet(s) arrived during current time slot will be served in the next time slot at the earliest.

4.1.2 Joint State Considering Channel Fading State and Channel Scheduling

We consider a homogenous system and analyze the performance of a particular user referred to as tagged user. Let us use random variable $m_n \in \{0,1\}$ to denote whether the channel is assigned to the tagged user or not at time slot n. $m_n = 1$ represents that the channel is assigned to the tagged user at time slot n and vice versa. The joint state at time slot n is denoted as

(m_n, c_n) and the transition probability can be written as follows

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{0\to 0} & \mathbf{J}_{0\to 1} \\ \mathbf{J}_{1\to 0} & \mathbf{J}_{1\to 1} \end{bmatrix}, \tag{4.1}$$

where $\mathbf{J}_{i\to j}(p,q)$ denotes the transition probability of the tagged SU channel assignment from state i to state j where $i, j \in \{0, 1\}$ and channel state goes from state p to state q where $0 \leq p, q \leq (Z-1)$. One can obtain the joint transition probability, $\mathbf{J}$ using the procedure mentioned in [18]. Let us construct a new state variable, $a_n = m_n c_n$ and state space, $\mathbf{\Theta} = \{a_n | 0 \leq a_n \leq Z-1\}$. According to [18], the transition probability matrix, $\mathbf{S}$ for the state space, $\mathbf{\Theta}$ can be expressed as follows

$$\mathbf{S}_{l\to\hat{l}} = \frac{\sum\limits_{x\in l} \sum\limits_{\hat{x}\in\hat{l}} \tau_x \mathbf{J}_{x\to\hat{x}}}{\sum\limits_{x\in l} \tau_x}, \tag{4.2}$$

where $\mathbf{S}_{l\to\hat{l}}$ denotes the transition probability from state l to state $\hat{l}$ and $l, \hat{l} \in \{0, 1, ..., Z-1\}$ and $\tau_{\mathbf{x}}$ denotes the steady state probability vector for the state space $\mathbf{\Theta}$ at state x. From a_n, one can obtain transmission rate of the tagged SU at time slot n, $R_n = r a_n$ where $0 \leq R_n \leq Y$ and Y is the maximum transmission rate given by $Y = r(Z-1)$.

4.2 Random Transmission Protocol

In our analysis, we consider a random transmission protocol that takes sensing errors into account. It is important to mention that in previous works [18], [69], [70], random transmission protocol was not considered. Let us use T to denote the transmission decision of CRN where $T \in\{0,1\}$. T takes value 1 when CRN decides to transmit otherwise T takes value 0. According to the random transmission protocol, CR BS transmits probabilistically based on the estimated channel occupancy, $\hat{O}_n$. The details of the protocol for downlink is described below:

- *Step 1*:At time slot n, CRN estimates PUs' channel occupancy, $\hat{O}_n$.
- *Step 2*: If $\hat{O}_n = 0$, CR BS decides to transmit with probability, P_1. On the other hand if $\hat{O}_n = 1$, CR BS decides to transmit with probability, P_2.[5]

With the traditional deterministic transmission protocol [18], [69]-[70], CR BS transmits with probability 1 if the channel is sensed as idle. If the channel is sensed as busy, CR BS does not attempt to transmit. So, it is obvious that the traditional deterministic protocol is a special case of the random transmission protocol when $P_1 = 0$ and $P_2 = 1$.

After a transmission is made by the CR BS, the transmitted packet(s) can be received successfully by the SU or they can collide with PUs' transmission based on the actual channel occupancy. If CR BS receives an acknowledgement from the corresponding SU, it discards the transmitted packet(s) from the buffer. Otherwise, it re-transmits the packet(s) in the

[5]The actual number of packets that will be transmitted depends on the buffer as well as channel fading state.

next transmission opportunity. So, there is no packet loss from the corresponding tagged SU's buffer due to the collision with PUs' transmission. Packets are lost only due to the buffer overflow. It is important to note that when there is a false alarm, by increasing P_1, CRN may improve SUs' transmission rate. On the other hand, by decreasing P_2, CRN can reduce the collision probability in presence of miss-detection. It will be shown in later in this chapter that our developed queuing analytic model can be used to determine these probabilities in order to maintain QoS requirements of both PUs and SUs for given other operating parameters.

4.3 Queuing Analytic Model

4.3.1 Joint State Considering Primary Users' Activity and Transmission Decision

According to the random transmission protocol, packet transmission depends on the transmission decision, T and actual PUs' channel occupancy, O. Therefore, we need to find joint transition probability of these random variables. Let us define the transition probability of the joint transmission decision and actual channel occupancy state, **R** as follows

$$\mathbf{R} = \begin{bmatrix} (T_{0\to 0}, O_{0\to 0}) & (T_{0\to 0}, O_{0\to 1}) & (T_{0\to 1}, O_{0\to 0}) & (T_{0\to 1}, O_{0\to 1}) \\ (T_{0\to 0}, O_{1\to 0}) & (T_{0\to 0}, O_{1\to 1}) & (T_{0\to 1}, O_{1\to 0}) & (T_{0\to 1}, O_{1\to 1}) \\ (T_{1\to 0}, O_{0\to 0}) & (T_{1\to 0}, O_{0\to 1}) & (T_{1\to 1}, O_{0\to 0}) & (T_{1\to 1}, O_{0\to 1}) \\ (T_{1\to 0}, O_{1\to 0}) & (T_{1\to 0}, O_{1\to 1}) & (T_{1\to 1}, O_{1\to 0}) & (T_{1\to 1}, O_{1\to 1}) \end{bmatrix}, \tag{4.3}$$

where $T_{m\to k}$ denotes the transition probability of transmission decision from state m to state k and $O_{a\to b}$ represents the transition probability of actual channel occupancy from state a to state b where m,k,a and $b \in \{0,1\}$. Mathematically, $(T_{m\to k}, O_{a\to b})$ can be written as follows

$$\begin{aligned} (T_{m\to k}, O_{a\to b}) \quad = \quad & [(1-P_1)^{(2-m-k)} P_1^{(m+k)} (1-P_d)^{(2-a-b)} P_f^{(a+b)} \\ & +(1-P_1)^{(1-m)} (1-P_2)^{(1-k)} P_1^{(m)} P_2^{(k)} \\ & (1-P_d)^{(1-a)} P_d^{(1-b)} (1-P_f)^{(b)} P_f^{(a)} \\ & +(1-P_2)^{(2-m-k)} P_2^{(m+k)} P_d^{(2-a-b)} (1-P_f)^{(a+b)} \\ & +(1-P_1)^{(1-k)} (1-P_2)^{(1-m)} P_1^{(k)} P_2^{(m)} \\ & (1-P_d)^{(1-b)} P_d^{(1-a)} (1-P_f)^{(a)} P_f^{(b)}] O_{a\to b}, \end{aligned} \tag{4.4}$$

where P_f, denotes the false alarm probability and P_d, denotes the miss-detection probability.

4.3.2 Markov Chain Analysis

We consider that state transition occurs at the slot boundary and state variables can take only a discrete value. So, we can model the system as a DTMC. State space of this DTMC for a finite size buffer can be denoted by $\mathbf{\Phi} \equiv \{b_n, R_n, T_n, O_n | 0 \leq b_n \leq Q; 0 \leq R_n \leq Y; 0 \leq T_n \leq 1; 0 \leq O_n \leq 1\}$, where b_n is the buffer length or number of packets in the buffer at a time slot n and Q in the maximum buffer size. After obtaining the transition probabilities of eq. (4.3), the state transition matrix for the DTMC can be written as in eq. (4.5) for an infinite buffer size. By making block of sub matrices of eq. (4.5) as a QBD process with level X where $X = \lfloor \frac{Q}{Y} \rfloor$, eq. (4.5) for the finite buffer case can be expressed as follows

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ X-1 \\ X \end{array} \begin{bmatrix} \mathbf{A} & \mathbf{C} & & & & & \\ \mathbf{D} & \mathbf{E} & \mathbf{F_0} & & & & \\ & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & & \\ & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0'} \\ & & & & & \mathbf{F_2'} & \mathbf{F_1'} \end{bmatrix}. \tag{4.6}$$

Let us use $\boldsymbol{\pi}_l$ to denote the vector of steady state probabilities of level l where $l = 0, 1, ..., X$. One can calculate the steady state probability vector $\boldsymbol{\pi} \equiv [\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, ..., \boldsymbol{\pi}_X]$ according to the procedure described in Appendix C.

$$\mathbf{P}_{\text{inf}} = \left[\begin{array}{c|ccccc|ccccc|ccc}
\mathbf{H}_{0}^{(0)} & \mathbf{H}_{1+}^{(0)} & \cdots & \mathbf{H}_{N+}^{(0)} & & & & & & & & & & \\
\hline
\mathbf{H}_{1-}^{(1)} & \mathbf{H}_{0}^{(1)} & \mathbf{H}_{1+}^{(1)} & \cdots & \mathbf{H}_{N+}^{(1)} & & & & & & & & & \\
\vdots & \vdots & & & \ddots & & & & & & & & & \\
\mathbf{H}_{(Y-N+1)-}^{(Y-N+1)} & \mathbf{H}_{(Y-N)-}^{(Y-N+1)} & \cdots & \mathbf{H}_{0} & \mathbf{H}_{(1)+}^{(Y-N+1)} & \mathbf{H}_{(N-1)+}^{(Y-N+1)} & \mathbf{H}^{N+} & & & & & & & \\
\vdots & \vdots & & & & \vdots & \vdots & \ddots & & & & & & \\
\mathbf{H}_{Y-}^{(Y)} & \mathbf{H}_{(Y-1)-}^{(Y)} & \cdots & \cdots & \mathbf{H}_{(1)-}^{(Y)} & \mathbf{H}_{0} & \mathbf{H}_{1+} & \cdots & \mathbf{H}_{N+} & & & & & \\
\hline
 & \mathbf{H}_{Y-} & \cdots & \cdots & \cdots & \mathbf{H}_{1-} & \mathbf{H}_{0} & \mathbf{H}_{1+} & \cdots & \mathbf{H}_{N+} & & & & \\
 & & \ddots & & & \vdots & \vdots & & & \ddots & & & & \\
 & & & \mathbf{H}_{Y-} & & \mathbf{H}_{(Y-N+1)-} & \mathbf{H}_{(Y-N)-} & \cdots & \cdots & \cdots & \mathbf{H}_{(N-1)+} & \mathbf{H}_{N+} & & \\
 & & & & \ddots & \vdots & \vdots & & & \ddots & \vdots & \vdots & \ddots & \\
 & & & & & \mathbf{H}_{Y-} & \mathbf{H}_{(Y-1)-} & \cdots & \cdots & \mathbf{H}_{1-} & \mathbf{H}_{0} & \mathbf{H}_{1+} & \cdots & \mathbf{H}_{N+} \\
\hline
 & & & & \ddots & & & & & \ddots & & & & \ddots
\end{array}\right], \tag{4.5}$$

4.4 Performance Analysis

Once the steady state probability vector, $\boldsymbol{\pi}$ of eq. (4.6) is obtained, one can measure different packet-level performance parameters. We are mainly interested to measure collision probability, packet loss probability due to the buffer overflow and average queuing delay for given values of P_1 and P_2 as well as other operating parameters.

4.4.1 Collision Probability

At time slot n, collision of a SU's transmission with PUs' transmission happens if the SU transmits one or more packets but the channel is actually occupied by a PU i.e., $O_n = 0$. CR BS makes a transmission decision i.e., $T_n = 1$, if buffer state of the SU, $b_n > 0$ and $R_n > 0$. Therefore, the collision probability corresponds to the summation of the steady state probabilities associated with the states having $T_n = 1$, $b_n > 0$, $R_n > 0$ and $O_n = 0$. Mathematically, the collision probability for the tagged SU can be calculated from the steady state probabilities as follows

$$P_{\text{tag,col}} = \sum_{b=1}^{Q} \sum_{R=1}^{Y} \boldsymbol{\pi}(b, R, 1, 0). \tag{4.7}$$

For a homogeneous system the collision probability due to the transmission of each SU will be identical to eq. (4.7). So the overall collision probability can be calculated as follows

$$P_{\text{col}} = P_{\text{tag,col}} \times K. \tag{4.8}$$

4.4.2 Packet Loss Probability

A packet of the tagged SU will be lost if it finds the buffer is full upon arrival. Considering steady state probabilities of those states that can lead to a buffer overflow and associated packet arrival probabilities, we can finally express the packet loss probability of the tagged SU as follows

$$P_{\text{loss}} = \sum_{b=(Q-N+1)}^{Q} \boldsymbol{\pi}(b, R, T, O)\mathbf{1} \times \sum_{i=(Q-b+1)}^{N} \beta_i, \tag{4.9}$$

where $\mathbf{1}$ is a column vector with length $4Z$.

4.4.3 Average Queuing Delay

Using the Little's law, the average queuing delay of a packet for the tagged SU can be written as follows

$$D_{\text{avg}} = \frac{\sum_{b=1}^{Q} p(b) \times b}{\beta' \times (1 - P_{\text{loss}})}, \tag{4.10}$$

where $p(b)$ denotes the probability of having b packets in the buffer. This probability, $p(b)$ can be obtained from the steady state probabilities as follows

$$p(b) = \boldsymbol{\pi}(b, R, T, O)\mathbf{1}, \qquad 1 \leq b \leq Q. \tag{4.11}$$

4.5 Numerical Results and Application

In this section, we present selected numerical examples to demonstrate the significance of our developed queuing analytic model. As a potential application of our developed queuing analytic model for a CAC decision in CRN, we provide an example. We also compare the performance of the considered random transmission protocol with the deterministic transmission protocol. All the presented numerical results are validated via computer simulation. In our numerical results, we assume that the maximum buffer size of each SU, Q=20, the number of channel states, Z=4 and PUs' activity, ρ=0.6. We also assume that the average packet arrival rate of each SU, β'=0.3, false alarm probability, P_f=0.3, miss-detection probability, P_d=0.3 and Nakagami fading parameter, $m = 1.1$.

4.5.1 Numerical Results

In Fig. 4.2, we have plotted collision probability for different number of SUs in the CRN. As the number of SUs increases, the collision probability increases. This can be explained as follows. As the number of SUs increases, the overall transmission probability increases due to multiuser diversity [18]. But due to the miss-detection, the increased transmission probability leads to a higher collision probability. Fig. 4.2 also shows that the effect of P_1 and P_2 of random transmission protocol on the collision probability. According to the random transmission protocol, the CR BS transmits with probability, P_1 if the channel is sensed as busy. Therefore, a decreased value of P_1 leads to a less aggressive transmission. As such a lower value of P_1 decreases the collision probability. Similarly, when there is a certain miss-detection, a decreased value of P_2 leads to a lower collision probability. So, a CRN can decrease the collision probability by reducing the values of P_1 and P_2.

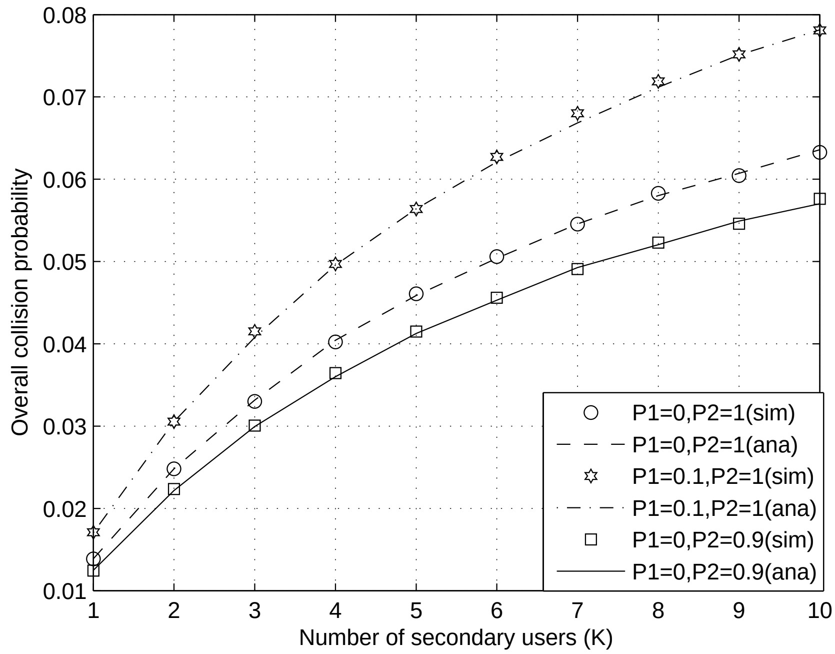

Figure 4.2: Effect of transmission probabilities, (P_1,P_2) and number of SUs, (K) on the collision probability.

Fig. 4.3 shows the effect of SUs on packet loss probability of the tagged SU. As the number of SUs increases, the transmission opportunity of each SU reduces. Therefore, the packet loss probability of each SU increases. In Fig. 4.4, we have plotted the effect of SUs on average queuing delay of the tagged SU which shows similar behavior as the packet loss probability plotted in Fig. 4.3. In Figs. 4.3 and 4.4, there are a negligible gap between the simulation and theoretical result. This can happen due to the numerical calculation of the steady state probabilities.

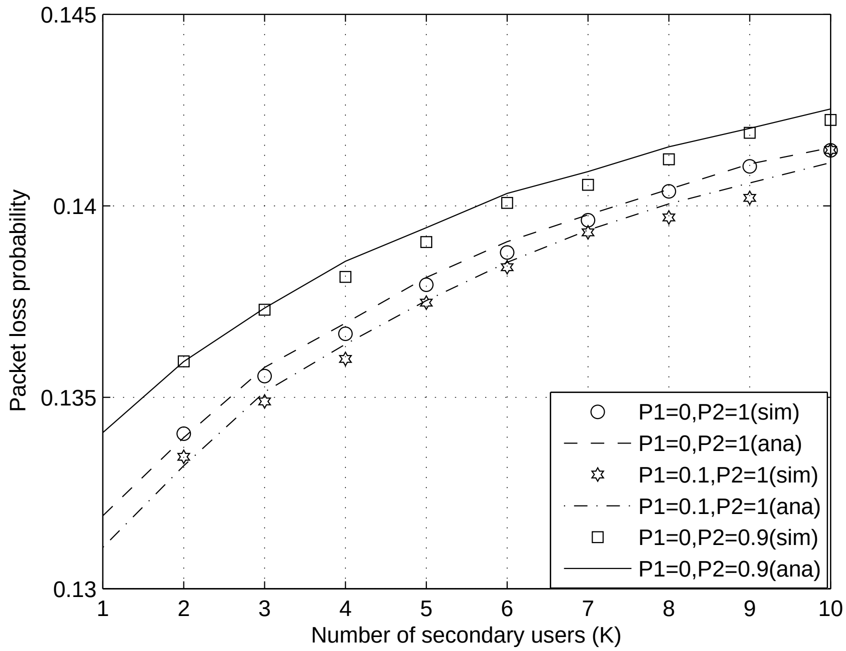

Figure 4.3: Effect of transmission probabilities, (P_1, P_2) and number of SUs, (K) on the packet loss probability.

Table 4.1: Transmission probabilities (P_1, P_2) vs. number of SUs for given QoS requirements ($p_{t,col} = 0.07$, $p_{t,ploss} = 0.14$ and $D_{t,avg} = 77$ (time slots))

(P_1, P_2)	$K_{p_{t,ploss}}$	$K_{D_{t,avg}}$	$K_{p_{t,col}}$	K_s
(0,0.9)	6	9	15	6
(0,1)	7	11	13	7
(0.1,1)	8	12	8	8

4.5.2 Application of Our Developed Analytic Model

In what follows, we explain an application of our developed queuing analytic model. In general, PUs can tolerate a certain probability of collision determined by the PUs' QoS requirement [71]. Let us assume that the PU system has specified a collision probability, $p_{t,col}$ of 0.07 as a threshold limit.

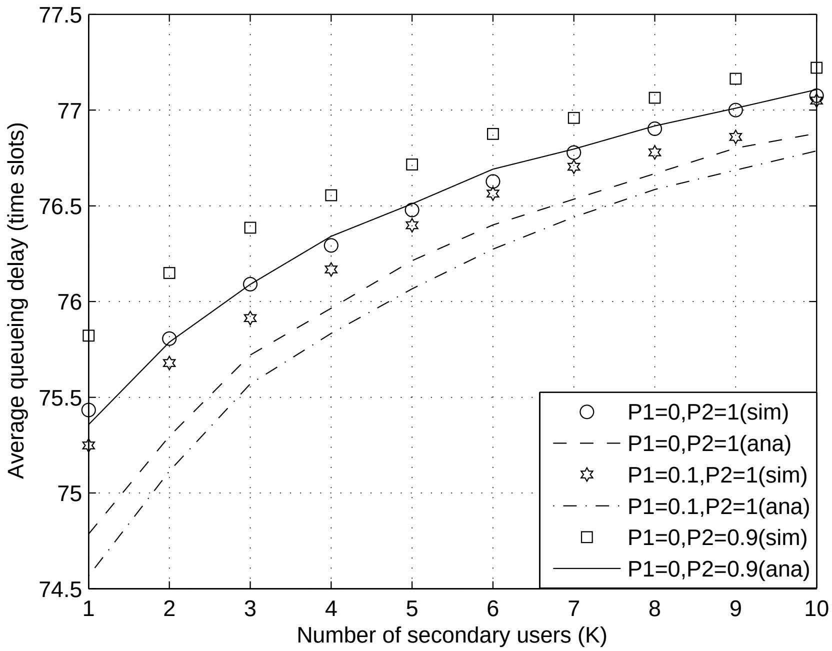

Figure 4.4: Effect of transmission probabilities, (P_1,P_2) and number of SUs, (K) on average queuing delay.

Let us consider that each SU has a target packet loss probability, $p_{t,ploss}$ of 0.14 and average queuing delay, $D_{t,avg}$ of 77 (time slots). Given other operating parameters, in order to maintain the above mentioned QoS requirements, the CRN can support a certain number of SUs for given values of P_1 and P_2.

Let us use $K_{p_{t,col}}$ to denote the maximum number of SUs that can be supported to maintain $p_{t,col}$ for given values of (P_1, P_2). From Fig. 4.2 one can obtain the value of $K_{p_{t,col}}$ for given values of (P_1, P_2). Similarly, let us use $K_{p_{t,ploss}}$ and $K_{D_{t,avg}}$ to denote, the maximum number of SUs that can be supported to satisfy $p_{t,ploss}$ and $D_{t,avg}$, respectively for particular values of (P_1, P_2). The values of $K_{p_{t,ploss}}$ and $K_{D_{t,avg}}$ can be obtained from Figs. 4.3 and 4.4, respectively. For example, different values of P_1, P_2 and corresponding $K_{p_{t,ploss}}$, $K_{D_{t,avg}}$, $K_{p_{t,col}}$ are listed in Table 4.1. For given values of P_1 and P_2, the maximum number of SUs that can be supported in order to satisfy all three QoS requirements is equal to the minimum of $K_{p_{t,ploss}}$, $K_{D_{t,avg}}$ and $K_{p_{t,col}}$ i.e., $K_s = \min(K_{p_{t,ploss}}, K_{D_{t,avg}}, K_{p_{t,col}})$ which is listed in the last column of Table 4.1.

From Table 4.1, it is obvious that CRN should choose the transmission probabilities of $P_1 = 0.1$ and $P_2 = 1$ as it can support the maximum number of SUs for the given QoS requirements and other operating parameters. From the column 5 of Table 4.1 it is clear that CRN can admit at most 8 SUs. It is important to note that the deterministic transmission protocol (i.e., $P_1 = 0$ and $P_2 = 1$) can support maximum 7 SUs. Interestingly, the random transmission protocol with probability $P_1 = 0.1$ and $P_2 = 1$ can support one more SU than the deterministic protocol.

Chapter 5

Conclusion and Future Works

5.1 Conclusion

In this thesis, we have made two major contributions. In the first part of this thesis, we have studied rate allocation mechanisms that allocate rate between two different classes of services namely BE and DS services of a particular SU. Specifically, we have formulated the optimal rate allocation mechanism as a MDP and the optimal rate allocation mechanism has been formulated using LP technique. We have also proposed a low-complexity suboptimal rate allocation mechanism. A queuing analytic model for the proposed suboptimal rate allocation mechanism has been developed in order to measure packet-level performance parameters. Selected numerical results have shown that the performance of the suboptimal rate mechanism is quite similar to the optimal rate allocation mechanism for considered operating parameters and QoS requirements. Our queuing analytic model for suboptimal mechanism is useful in making a CAC decision in CRN for specific QoS requirements for two different classes of services.

In the second part, we have developed a queuing analytic model for a random transmission protocol that incorporated sensing errors into account in order to measure packet-level performance parameters as well as collision probability with PUs' transmission. The analytic model is useful for a CAC decision in CRN for specific QoS requirements in presence of sensing errors. We have also compared the performance of the random transmission protocol with the deterministic transmission protocol. Presented numerical results have shown that by properly selecting the transmission probabilities of the random transmission protocol, CRN can support a higher number of SUs than the deterministic transmission protocol while satisfying specified QoS requirements. In particular, for the case presented in this thesis, the random transmission protocol can support about 15% more SUs compared to the deterministic transmission protocol.

5.2 Future Works

This thesis has opened up some interesting research problems that may be pursued as future works.

- The analytic model presented in Chapter 3 is based on the assumption that the channels are perfectly sensed. It is interesting to extend the model with imperfect channel sensing.

- The analytic model presented in Chapter 4 is developed for a single channel CRN. Further research is needed to develop analytic model for a multi-channel CRN.

– In this thesis, we have not prioritized SUs at the user level. It is interesting to develop rate allocation mechanism for prioritized SUs where each SU supports different classes of services.

Bibliography

[1] B. Wang and K. J. R. Liu, "Advances in cognitive radio networks: a survey", *IEEE J. Sel. Topics in Signal Process.*, vol. 5, pp. 5-23, Feb. 2011. → pages 1

[2] Y. C. Liang, K. C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview", *IEEE Trans. on Veh. Tech.*, vol. 60, pp. 3386-3407, Sep. 2011. → pages x, 1, 4

[3] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks", *IEEE Commun. Mag.*, vol. 46, pp. 40-48, Apr. 2008. → pages 2

[4] M. T. Mushtaq, M. S. Khan, M. R. Naqvi, R. D. Khan, M. A. Khan, and O. F. Koudelka1, "Cognitive radios and cognitive networks: A short introduction", *J. Basic Applied Scientific Research*, pp. 56-65, 2013. → pages 2

[5] V. Chakravarthy, X. Li, Z. Wu, M. Temple, F. Garber, R. Kannan, and A. Vasilakos, "Novel overlay/underlay cognitive radio wave forms using sd-smse framework to enhance spectrum efficiency- part I: The-

oretical framework and analysis in awgn channel", *IEEE Trans. on Commun.*, vol. 57, pp. 3794-3804, Dec. 2009. → pages 3

[6] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks", *IEEE Trans. on Wireless Commun.*, vol. 7, pp. 5306-5315, Dec. 2008. → pages 3, 7, 24

[7] T. C. Clancy, "Dynamic spectrum access using the interference temperature model", *Annals of Telecommunications, Springer*, vol. 64, no. 7, pp. 573-592, 2009. → pages 3

[8] X. Zhang and H. Su, "CREAM-MAC: cognitive radio-enabled multi-channel mac protocol over dynamic spectrum access networks", *IEEE J. Sel. Topics in Signal Process.*, vol. 5, pp. 110-123, Feb. 2011. → pages 5

[9] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal", *IEEE Pers. Commun.*, vol. 6, pp. 13-18, Aug. 1999. → pages 5

[10] ET Docket No. 03-108, 20 FCC Rcd 5486, "Cognitive radio report and order", Mar. 2005. Summary available at `http://edocket.access.gpo.gov/2005/05-8808.htm`. → pages 5

[11] 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 Commun. Mag.*, vol. 46, pp. 72-79, Jul. 2008. → pages 5

[12] S. Haykin, "Cognitive radio: Brain-empowered wireless communications", *IEEE J. Sel. Areas in Commun.*, vol. 23, pp. 201-220, Feb. 2005. → pages 5

[13] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/ dynamic spectrum access/cognitive radio wireless networks: A survey,", *Comput. Netw.*, vol. 50, pp. 2127-2159, May 2006. → pages 5

[14]] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks,", *Ad Hoc Netw.*, vol. 7, pp. 810-836, Jul. 2009. → pages 6

[15] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio", *IEEE Trans. on Inform. Theory*, vol. 52, pp. 1813-1827, May 2006. → pages 6

[16] S. A. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity", *IEEE J. Select. Areas in Commun.*, vol. 25, pp. 529-537, Apr. 2007. → pages 6

[17] S. Srinivasa and S. A. Jafar, "The throughput potential of cognitive radio: a theoretical perspective", *IEEE Commun. Mag.*, vol. 45, pp. 73-79, May 2007. → pages 6

[18] M. M. Rashid, Md. J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum scheduling for multiuser cognitive radio: A queueing analysis", *IEEE Trans. on Wireless Commun.*, Vol. 8, pp. 5259-5269, Oct. 2009. → pages 6, 14, 15, 16, 24, 29, 43, 59, 63, 64, 71

[19] S. Keshavamurthy and K. Chandra, "Multiplexing analysis for dynamic spectrum access", in Proc. of *IEEE Military Commun. Conf. (MILCOM'06)*, Washington, DC, USA, pp. 1-7, Oct. 2006. → pages 6, 24

[20] S. Tang, and B. L. Mark, "Performance analysis of a wireless network with opportunistic spectrum sharing", in Proc. of *IEEE Global Commun. Conf. (Globecom'07)*, Washington, DC, USA, pp. 4636-4640, Nov. 2007. → pages 6, 24

[21] C. Zhang, X. Wang, and J. Li, "Cooperative cognitive radio with priority queueing analysis", in Proc. of *IEEE Int. Conf. on Commun. (ICC'09)*, Dresden, Germany, Jun. 2009, pp. 1-5. → pages 6, 24

[22] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisioning over cognitive radio wireless networks", *IEEE J. Sel. Areas in Commun.*, vol. 26, pp. 118-129, Jan. 2008. → pages 7, 24

[23] L. T. Tan, and L. B. Le, "Channel assignment for throughput maximization in cognitive radio networks", in Proc. of *IEEE Wireless Commun. and Networking Conf. (WCNC'12)*, China, pp. 1427-1431, Apr. 2012. → pages 7, 24

[24] N. Nie and C. Comaniciu, "Adaptive channel allocation spectrum etiquette for cognitive radio networks", in Proc. of *IEEE Dynamic Spectrum Access Networks (DySPAN'05)*, Baltimore, MD, USA, pp. 269-278, Nov. 2005. → pages 7, 24

[25] P. Frossard, C. W. Chen, C. J. Sreenan, K. P. Subbalakshmi, D.O. Wu, and Q. Zhang, "Guest Editorial Cross-layer Optimized Wireless Multimedia Communication", *IEEE J. Sel. Areas in Commun.*, vol. 25, pp. 641-644, May 2007. → pages 7

[26] Md. J. Hossain, "Some cross-layer design and performance issues in wireless networks", Master's thesis, available online *https://circle.ubc.ca/handle/2429/31332*, Oct. 2007. → pages 7

[27] K. B. Johnsson and D. C. Cox, "QoS scheduling of mixed priority non-real-time traffic", in Proc. of *IEEE Veh. Tech. Conf. (VTC'spring)*, vol. 4, Rhodes, Greece, May 2001, pp. 2645-2649. → pages 7

[28] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks", *IEEE Commun. Mag.*, vol. 41, pp. 74-80, Oct. 2007. → pages 7

[29] P. Frossard, C. Chen, C. Sreenan, K. Subbalakshmi, D. Wu, and Q. Zhang, "Guest editorial cross-layer optimized wireless multimedia communications", *IEEE J. Sel. Areas in Commun.*, vol. 24, pp. 641-644, May 2007. → pages 8

[30] D. I. Kim, L. B. Le, and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment", *IEEE Trans. Wireless Commun.*, vol. 7, pp. 55175527, Dec. 2008. → pages 8

[31] A.T. Hoang and Y.C. Liang, "Maximizing Spectrum Utilization of Cognitive Radio Networks Using Channel Allocation and Power Con-

trol", in Proc. *IEEE Veh. Tech. Conf. (VTC 06 Fall)*, Sept. 2006. → pages 8

[32] A. T. Hoang, Y.-C. Liang, and M. H. Islam, 'Power control and channel allocation in cognitive radio networks with primary users' cooperation", *IEEE Trans. on Mobile Computing*, vol. 9, pp. 348-360, 2010. → pages 8

[33] T. A. Khasib and L. Lampe, "Admission control and resource optimization for multiple-user OFDMA cognitive radio systems", available online *http://www.ece.ubc.ca/~lampe/Preprints/Admission_Cognitive.pdf*, 2012. → pages 8

[34] D. Pacheco-Paramo, V. Pla, and J. Martinez-Bauset, "Optimal admission control in cognitive radio networks", in Proc. of *4th Int. Conf. on Cognitive Radio Oriented Wireless Networks and Commun. (CROWNCOM'09)*, Hannover, Germany, pp. 281-287, Jun. 2009. → pages 8

[35] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications", *IEEE Commun. Surveys and Tutorials*, vol. 11, pp. 116-130, 2009. → pages 8, 58

[36] K. Wu, W. Wang, H. Luo, G. Yu, and Z. Zhang, "Optimal resource allocation for cognitive radio networks with imperfect spectrum sensing", in Proc. of *IEEE Veh. Tech. Conf. (VTC'10)*, Taiwan, May. 2010, pp. 1-4. → pages 8, 31

[37] S. Tang, B.L. Mark, "Modeling and analysis of opportunistic spectrum sharing with unreliable spectrum sensing," *IEEE Trans. on Wireless Commum.*, vol. 8, pp. 1934-1943, Apr. 2009. → pages 8

[38] S. Haykin, D. Thomson, and J. Reed, "Spectrum sensing for cognitive radio", *Proc. of the IEEE*, vol. 97, pp. 849-877, May 2009. → pages 8

[39] L. C. Horvath and J. Bito, "Primary and secondary user activity models for cognitive wireless network", in Proc. of *International Conference on Telecommunications (ConTEL'11)*, Graz, Austria, Jun. 2011, pp. 301-306. → pages 13

[40] Q. Zhao, L. Tong, A. Swami, and Y Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework", *IEEE J. Sel. Areas in Commun.*, vol. 25, pp. 589 - 600, Apr. 2007. → pages 13

[41] S. Akin and M. C. Gursoy, "Performance analysis of cognitive radio systems under QoS constraints and channel uncertainty", *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 2883-2895, Sep. 2011. → pages 13

[42] W. Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks", *IEEE Trans. on Wireless Commun.*, vol. 7, pp. 3845-3857, Oct. 2008. → pages 13

[43] M. M. Rashid, M. J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum access in cognitive radio networks: A queuing analytic model and admission controller design", in Proc. of *IEEE*

Global Commun. Conf. (Globecom'07), Washington, DC, USA, pp. 4647-4652, Nov. 2007. → pages 14

[44] V. K. Tumuluru, P. Wang, and D. Niyato, "A novel spectrum-scheduling scheme for multichannel cognitive radio network and performance analysis", *IEEE Trans. on Vech. Tech.*, vol. 60, pp. 1849-1858, May 2011. → pages 14

[45] A. goldsmith, "Wireless Communications," 1st Ed., Cambridge university press, 2005, pp. 79-82. → pages 15

[46] H. S. Wang and N. Moayeri, "Finite-state Markov channel a useful model for radio communication channels", *IEEE Trans. on Vech. Tech.*, vol. 44, pp. 163-171, Feb. 1995. → pages 15

[47] Q. Liu, S. Zhou, and G. B. Giannakis, "Queuing with adaptive modulation and coding over wireless links: cross-layer analysis and design", *IEEE Trans. on Wireless Commun.*, vol. 4, pp. 1142-1153, May 2005. → pages 15

[48] A. J. Goldsmith and S. G. Chua, "Adaptive coded modulation for fading channels", *IEEE Trans. on Commun.*, vol. 46, pp. 595-602, May 1998. → pages 15

[49] R. Urgaonkar and M. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks", *IEEE Trans. on Mobile Computing.*, vol. 8, pp. 766-777, Jun. 2009. → pages 16

[50] Z. Liang, S. Feng, D. Zhao, and X. Shen, "Delay performance analysis

for supporting real-time trafc in a cognitive radio sensor network", *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 325-335, Jan. 2011. → pages 16

[51] H. Lee and D. H. Cho, "VoIP capacity analysis in cognitive radio system", *IEEE Commun. Letters*, vol. 13, pp. 393-395, Jun. 2009. → pages 16

[52] J.S.H. Leeuwaarden and E.M.M. Winands, "Quasi-birth-death processes with an explicit rate matrix", *Stochastic Models*, Vol. 22, No. 1, pp. 77 98, 2006. → pages 18

[53] J. P. Kharoufeh, "Level-dependent quasi-birth-and-death processes", available online *http://pitt.edu/~jkharouf/Papers/LDQBD_Kharoufeh_v1.pdf*. → pages 18

[54] G. Bolch, S. Greiner, H. de Meer and K. S. Trivedi, "Queueing Networks and Markov Chains," 2nd Ed., Hoboken, Ed. New Jersey: John Wiley & Sons, Inc., 2006, pp. 127-140. → pages 19, 20, 21, 108

[55] M. L. Puterman, "Markov Decision Processes: Discrete Stochastic Dynamic Programming," 1st Ed., John Wiley & Sons, Inc., 1995, pp. 391-397. → pages x, 21, 36, 38

[56] M. H. Ngo, "Cross-layer adaptive transmission scheduling in wireless networks", PhD thesis, available online *https://circle.ubc.ca/handle/2429/1626?show=full*, Nov. 2003. → pages 22, 23

[57] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, "Optimal and

suboptimal packet scheduling over correlated time varying flat fading channels", *IEEE Trans. on Wireless Commun.*, vol. 5, pp. 446-456, Feb. 2006. → pages 23, 35, 37, 38

[58] V. K. Tumuluru, P. Wang, D. Niyato, and W. Song, "Performance analysis of cognitive radio spectrum access with prioritized traffic", *IEEE Trans. on Veh. Tech.*, vol. 61, pp. 1895-1906, May 2012. → pages 25

[59] Y. Lee, C. G. Park, and D. B. Sim, "Cognitive radio spectrum access with prioritized secondary users", *Appl. Math Inf. Sci*, vol. 6, pp. 595-601, Jan. 2012. → pages 25

[60] C. Ghosh , S. Chen , D. P. Agarwal, and A. M. Wyglinski, "Priority-based spectrum allocation for cognitive radio networks employing NC-OFDM transmission", in Proc. of *IEEE Military Commun. Conf. (MILCOM'09)*, Boston, MA, USA, Oct. 2009, pp.1 -5. → pages 25

[61] I. A. M. Balapuwaduge, "Performance evaluation of channel aggregation strategies in cognitive radio networks with queues", Master's thesis, available online *http://brage.bibsys.no/hia/bitstream/URN:NBN:no-bibsys_brage_33322/1/masteroppgave.pdf*, May 2012. → pages 25

[62] Md. J. Hossain, D. V. Djonin, and V. K. Bhargava, "Delay limited optimal and suboptimal power and bit loading algorithms for OFDM systems over correlated fading channels", in Proc. of *IEEE Global*

Commun. Conf. (GLOBECOM'05), St. Louis, MO, USA, Dec. 2005, pp. 2787-2792. → pages 36, 37

[63] K. Murphy, available online *http://www.cs.ubc.ca/~murphyk/Software/MDP/mdp.html*. → pages 37

[64] W. Liu, W. Lou, and Y. Fang, "An efficient quality of service routing algorithm for delay-sensitive applications", *Computer Networks*, vol. 47, pp.87-104, 2005. → pages 42

[65] A. Abdrabou, and W. Zhuang, "Stochastic delay guarantees and statistical call admission control for IEEE 802.11 single-hop ad hoc networks", *IEEE Trans. on Wireless Commun.*, vol. 7, pp. 3972-3981, Oct. 2008. → pages 42

[66] T. Yucek, and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys and Tutorials*, Vol. 11, pp. 116-130, First Quarter, 2009. → pages 58

[67] Y. Zeng , Y. C. Liang , A. T. Hoang , R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions", *EURASIP J. on Advances in Signal Processing*, available online *http://asp.eurasipjournals.com/content/2010/1/381465*, Jan. 2010, → pages 58

[68] P. Wang, L. Xiao, S. Zhou, and J. Wang, "Optimization of detection time for channel efficiency in cognitive radio systems", in Proc. *c. IEEE Wireless Commun. and Networking Conf. (WCNC'2007)*, Hong Kong, Mar. 2007, pp. 111-115. → pages 58

[69] M. Zhang, S. Jiang, G. Wei, and H. Wang, "Performance analysis of the cognitive radio network with a call level queue for secondary users", in Proc. of *5th Int. Conf. on Wireless Commun., Network. and Mobile Computing*, China, Sept. 2009, pp. 1-4. → pages 59, 64

[70] J. Wang, A. Huang, L. Cai, and W. Wang, "On the queue dynamics of multi-user multi-channel cognitive radio networks", *IEEE Trans. on Veh. Technol.*, Vol. 62, pp. 1314-1328, Mar. 2013. → pages 59, 64

[71] Y. Pei, A. T. Hoang, and Y. Liang, "Sensing-throughput tradeoff in cognitive radio networks: How frequently should spectrum sensing be carried out?, " in Proc. of *IEEE Int. Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC'07)*, Singapore, Sep. 2007, pp. 1-5. → pages 73

Appendices

Appendix A

Let us use W to denote the maximum number of packet(s) to be transmitted from BE service buffer when action $a^{(t)}$ is taken i.e., $W = Y - a^{(t)}$ where $a^{(t)} \in \mathcal{A}$ and Y is the maximum transmission rate. Let us consider a set of sub matrices $\mathbf{V}^{(x)}(0 \leq x \leq \lfloor \frac{W}{r} \rfloor)$ and it can be written as follows

$$\mathbf{V}_{i,j}^{(x)} = \begin{cases} \mathbf{S}_{i,j} & \text{if} \quad i = x \\ 0 & \text{if} \quad i \neq x \end{cases}, \tag{1}$$

where $0 \leq i, j \leq \lfloor \frac{W}{r} \rfloor$ and $\mathbf{S}_{i,j}$ denotes the transition probability of transmission rate from state i to state j. Eq. (1) can be rewrite as follows

$$\mathbf{V}_{i,j}^{(x)} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ S_{x,0} & S_{x,1} & \cdots & S_{x,Y} \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix}. \tag{2}$$

Please note that for any value e.g., $\mathbf{V}_{k,j}^{(e)} = 0$ if $e \notin x$, $\alpha_p = 0$ if $p \notin M$ and $\beta_q = 0$ if $q \notin N$. The transition probability matrix for optimal rate allocation mechanism when action $a^{(t)}$ is taken can be expressed as eq. (3).

where $\mathbf{B}_{i-}^{(j)}$ represents the net decrement of i packet(s) from the BE buffer when the buffer had j packet(s), $\mathbf{B}_{i+}^{(j)}$ represents the net increment of i packet(s) from the BE buffer when the buffer had j packet(s), $\mathbf{B}_{i-}$ represents the net decrement of i packet(s) from the BE buffer and $\mathbf{B}_{i+}$ represents the net increment of i packet(s) from the BE buffer. For finite buffer sizes, eq. (3) can be expressed as follows

$$\mathbf{P}^{(o,a^{(t)})} = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ XX-1 \\ XX \end{matrix} \begin{bmatrix} \mathbf{A} & \mathbf{C} & & & & & \\ \mathbf{D} & \mathbf{E} & \mathbf{F_0} & & & & \\ & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & & \\ & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \mathbf{F_2} & \mathbf{F_1} & \mathbf{F_0'} \\ & & & & & \mathbf{F_2'} & \mathbf{F_1'} \end{bmatrix}, \qquad (4)$$

where $XX = \lfloor \frac{Q^{(b)}}{W} \rfloor$.

$$\mathbf{B}^{(o,a^{(t)})} = \begin{bmatrix} \mathbf{H}_0^{(0)} & \mathbf{H}_{1+}^{(0)} & \cdots & \mathbf{H}_{N+}^{(0)} & & \\ \mathbf{H}_{1-}^{(1)} & \mathbf{H}_0^{(1)} & \mathbf{H}_{1+}^{(1)} & \cdots & \mathbf{H}_{N+}^{(1)} & \\ \vdots & & \ddots & & \ddots & \\ \mathbf{H}_{a^{(t)}-}^{(a^{(t)})} & \cdots & \cdots & \cdots & \mathbf{H}_{(N)+}^{(a^{(t)})} & \\ \ddots & & \ddots & & & \vdots \\ & \mathbf{H}_{a^{(t)}-}^{(Q^{(d)}-q)} & \cdots & \cdots & \cdots & \mathbf{H'}_{(q-1)+}^{(Q^{(d)}-q)} \\ & \ddots & \ddots & & & \vdots \\ & & \mathbf{H}_{a^{(t)}-}^{(Q^{(d)})} & \cdots & \cdots & \mathbf{H'}_{(0)}^{(Q^{(d)})} \end{bmatrix}, \qquad (5)$$

$$
\mathbf{P}_{\mathbf{inf}}^{(\mathbf{o},\mathbf{a}^{(\mathbf{t})})} = \left[\begin{array}{c|ccccc|ccccc|cc}
\mathbf{B}_{0}^{(0)} & \mathbf{B}_{1+}^{(0)} & \cdots & \mathbf{B}_{M+}^{(0)} & & & & & & & & & \\
\hline
\mathbf{B}_{1-}^{(1)} & \mathbf{B}_{0}^{(1)} & \mathbf{B}_{1+}^{(1)} & \cdots & \mathbf{B}_{M+}^{(1)} & & & & & & & & \\
\vdots & \vdots & & & \ddots & & & & & & & & \\
\mathbf{B}^{(W-M+1)}_{(W-M+1)-} & \mathbf{B}^{(W-M+1)}_{(W-M)-} & \cdots & \mathbf{B}_{0} & \mathbf{B}^{(W-M+1)}_{(1)+} & \mathbf{B}^{(W-M+1)}_{(M-1)+} & \mathbf{B}_{M+} & & & & & & \\
\vdots & \vdots & & & & \vdots & \vdots & \ddots & & & & & \\
\mathbf{B}^{(W)}_{W-} & \mathbf{B}^{(W)}_{(W-1)-} & \cdots & \cdots & \mathbf{B}^{(W)}_{(1)-} & \mathbf{B}_{0} & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & & & \\
\hline
 & \mathbf{B}_{W-} & \cdots & \cdots & \cdots & \mathbf{B}_{1-} & \mathbf{B}_{0} & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & & \\
 & & \ddots & & & \vdots & \vdots & & & \ddots & & & \\
 & & & \mathbf{B}_{W-} & & \mathbf{B}_{(W-M+1)-} & \mathbf{B}_{(W-M)-} & \cdots & \cdots & \cdots & \mathbf{B}_{(M-1)+} & \mathbf{B}_{M+} & \\
 & & & & \ddots & \vdots & \vdots & & & \ddots & \vdots & \vdots & \\
 & & & & & \mathbf{B}_{W-} & \mathbf{B}_{(W-1)-} & \cdots & \cdots & \mathbf{B}_{1-} & \mathbf{B}_{0} & \mathbf{B}_{1+} & \cdots \\
\hline
 & & & & \ddots & & & & & \ddots & & &
\end{array}\right], \tag{3}
$$

where $\mathbf{H}_{i-}^{(j)}$ represents net decrement of i packet(s) from the DS buffer when the buffer had j packet(s), $\mathbf{H}_{i+}^{(j)}$ represents net increment of i packet(s) from the DS buffer when the buffer had j packet(s). Using eqs. (1) and (5), the inner sub matrices of eq. (4) can be written as follows

$$\mathbf{B}_{L-} = \begin{cases} \mathbf{H}_l^{(e)} &= \beta_{|a^{(t)}-l|} \sum\limits_{\substack{0 \le j \le M \\ j+L+a^{(t)} \mod r=0}} \alpha_j \mathbf{V}^{(j+L+a^{(t)})} \\ \mathbf{H}'_q &= (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \sum\limits_{\substack{0 \le j \le M \\ j+L+a^{(t)} \mod r=0}} \alpha_j \mathbf{V}^{(j+L+a^{(t)})}, \end{cases} \tag{6}$$

$$\mathbf{B}_{L} = \begin{cases} \mathbf{H}_l^{(e)} &= \beta_{|a^{(t)}-l|} \sum\limits_{\substack{0 \le j \le M-L \\ j+a^{(t)} \mod r=0}} \alpha_{j+L} \mathbf{V}^{(j+a^{(t)})} \\ \mathbf{H}'_q &= (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \sum\limits_{\substack{0 \le j \le M-L \\ j+a^{(t)} \mod r=0}} \alpha_{j+L} \mathbf{V}^{(j+a^{(t)})}, \end{cases} \tag{7}$$

$$\mathbf{B}_{L}^{(0)} = \begin{cases} \text{if } e = a^{(t)} & \\ \mathbf{H}_l^{(e)} &= \beta_{|a^{(t)}-l|} \alpha_L \sum\limits_{\lceil \frac{a^{(t)}}{r} \rceil \le j \le N_p Z} \mathbf{V}^{(j)} \\ \mathbf{H}'_q &= (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \alpha_L \sum\limits_{\lceil \frac{a^{(t)}}{r} \rceil \le j \le N_p Z} \mathbf{V}^{(j)} \\ \text{if } e > a^{(t)} & \\ \mathbf{H}_l^{(e)} &= \beta_{|a^{(t)}-l|} \alpha_L \mathbf{V}^{(a^{(t)})} \\ \mathbf{H}'_q &= (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \alpha_L \mathbf{V}^{(a^{(t)})}, \end{cases} \tag{8}$$

$$
\mathbf{B}_L^{(E)} = \begin{cases}
\text{if } e = a^{(t)} & \\
\mathbf{H}_l^{(e)} & = \beta_{|a^{(t)}-l|}\{ \sum_{\substack{0\le j\le E \\ j+a^{(t)} \mod r=0}} \alpha_{j+L}\mathbf{V}^{(j+a^{(t)})} \\
& + \alpha_{L+E} \sum_{\lceil \frac{a^{(t)}+E+1}{r}\rceil \le k \le N_p Z} \mathbf{V}^{(k)}\} \\
\mathbf{H}_q' & = (1 - \sum_{0\le j\le a^{(t)}+q-1} \beta_j)\{ \sum_{\substack{0\le j\le E \\ j+a^{(t)} \mod r=0}} \alpha_{j+L}\mathbf{V}^{(j+a^{(t)})} \\
& + \alpha_{L+E} \sum_{\lceil \frac{a^{(t)}+E+1}{r}\rceil \le k \le N_p Z} \mathbf{V}^{(k)}\} \\
\text{if } e > a^{(t)} & \\
\mathbf{H}_l^{(e)} & = \beta_{|a^{(t)}-l|} \sum_{\substack{0\le j\le E \\ j+a^{(t)} \mod r=0}} \alpha_{j+L}\mathbf{V}^{(j+a^{(t)})} \\
\mathbf{H}_q' & = (1 - \sum_{0\le j\le a^{(t)}+q-1} \beta_j) \sum_{\substack{0\le j\le E \\ j+a^{(t)} \mod r=0}} \alpha_{j+L}\mathbf{V}^{(j+a^{(t)})}
\end{cases} \tag{9}
$$

$$
\mathbf{B}_{L-}^{(E)} = \begin{cases}
\text{if } e = a^{(t)} \\
\mathbf{H}_l^{(e)} = \beta_{|a^{(t)}-l|}\{ \sum\limits_{\substack{0 \le j \le E-L \\ j+a^{(t)}+L \mod r=0}} \alpha_j \mathbf{V}^{(j+a^{(t)}+L)} \\
\qquad + \alpha_{E-L} \sum\limits_{\lceil \frac{a^{(t)}+E+1}{r} \rceil \le k \le N_p Z} \mathbf{V}^{(k)} \} \\
\mathbf{H}'_q = (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \{ \sum\limits_{\substack{0 \le j \le E-L \\ j+a^{(t)}+L \mod r=0}} \alpha_j \mathbf{V}^{(j+a^{(t)}+L)} \\
\qquad + \alpha_{E-L} \sum\limits_{\lceil \frac{a^{(t)}+E+1}{r} \rceil \le k \le N_p Z} \mathbf{V}^{(k)} \} \\
\text{if } e > a^{(t)} \\
\mathbf{H}_l^{(e)} = \beta_{|a^{(t)}-l|} \sum\limits_{\substack{0 \le j \le E-L \\ j+a^{(t)}+L \mod r=0}} \alpha_j \mathbf{V}^{(j+a^{(t)}+L)} \\
\mathbf{H}'_q = (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \sum\limits_{\substack{0 \le j \le E-L \\ j+a^{(t)}+L \mod r=0}} \alpha_j \mathbf{V}^{(j+a^{(t)}+L)}
\end{cases} \tag{10}
$$

$$
\mathbf{B}'_k = \begin{cases}
\mathbf{H}_l^{(e)} = \beta_{|a^{(t)}-l|} \sum\limits_{k \le j \le M} \sum\limits_{\substack{k \le h \le j \\ h+a^{(t)} \mod r=0}} \alpha_j \mathbf{V}^{(h+a^{(t)})} \\
\mathbf{H}'_q = (1 - \sum\limits_{0 \le j \le a^{(t)}+q-1} \beta_j) \sum\limits_{k \le j \le M} \sum\limits_{\substack{k \le h \le j \\ h+a^{(t)} \mod r=0}} \alpha_j \mathbf{V}^{(h+a^{(t)})}
\end{cases} \tag{11}
$$

$$
\mathbf{F}'_1 = \begin{bmatrix}
\mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & \\
\mathbf{B}_{1-} & \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & \\
\vdots & & \ddots & & \ddots & \\
\mathbf{B}_{(W-M+XX')-} & \cdots & \cdots & \cdots & & \mathbf{B}'_{(M-1)+} \\
\vdots & & \ddots & & & \vdots \\
\mathbf{B}_{W-} & \cdots & \cdots & \cdots & \cdots & \mathbf{B}'_{(XX'-1)+} \\
\ddots & & \ddots & & & \vdots \\
 & & \mathbf{B}_{W-} & \cdots & \cdots & \mathbf{B}'_{(0)}
\end{bmatrix}, , \tag{12}
$$

$$\mathbf{F}_0' = \begin{bmatrix} \mathbf{F_0} & \mathbf{0}_{((Q^{(d)}+1)\times WZ)\times((Q^{(d)}+1)\times XX'Z)} \end{bmatrix}, \tag{13}$$

$$\mathbf{F}_2' = \begin{bmatrix} \mathbf{F_2} \\ \mathbf{0}_{((Q^{(d)}+1)\times XX'Z)\times((Q^{(d)}+1)\times WZ)} \end{bmatrix}, \tag{14}$$

where $0 \le q \le N - a^{(t)} - 1$, $0 \le k \le M - 1$, $0 \le L \le M$, $1 \le L- \le W$, $-a^{(t)} \le l \le -a^{(t)} + N$, $1 \le E \le Q^{(b)}$, $0 \le e \le Q^{(d)}$ and $XX' = Q^{(b)} - \lfloor \frac{Q^{(b)}}{W} \rfloor \times W$. After obtaining transition probability of all actions of eq. (4), one can get the transition probability matrix for the optimal rate allocation mechanism, $P^{(o)}$ as follows

$$\mathbf{P}^{(o)} = \begin{bmatrix} \mathbf{P}^{(o,a^{(0)})} \\ \mathbf{P}^{(o,a^{(r)})} \\ \mathbf{P}^{(o,a^{(2r)})} \\ \vdots \\ \mathbf{P}^{(o,a^{(Y)})} \end{bmatrix}. \tag{15}$$

Appendix B

Let us consider a set of sub matrices $\mathbf{U}^{(x)}(0 \le x \le N_pZ)$ and it can be written as follows

$$\mathbf{U}^{(x)}_{k,j} = \begin{cases} \mathbf{S}_{k,j} & \text{if} \quad k = x \\ 0 & \text{if} \quad k \neq x \end{cases}, \tag{16}$$

where $0 \le k, j \le N_pZ$ and $\mathbf{S}_{k,j}$ denotes the transition probability from state k to state j. Please note that for any value e.g., $\mathbf{U}^{(e)}_{k,j} = 0$ if $e \notin x$, $\alpha_p = 0$ if $p \notin M$ and $\beta_q = 0$ if $q \notin N$.

$$\mathbf{B} = \begin{bmatrix} \mathbf{H}^{(0)}_{0} & \mathbf{H}^{(0)}_{1+} & \cdots & \mathbf{H}^{(0)}_{N+} & & \\ \mathbf{H}^{(1)}_{1-} & \mathbf{H}^{(1)}_{0} & \mathbf{H}^{(1)}_{1+} & \cdots & \mathbf{H}^{(1)}_{N+} & \\ \vdots & & \ddots & & \ddots & \\ \mathbf{H}^{(Y)}_{Y-} & \cdots & \cdots & \cdots & \mathbf{H}^{(Y)}_{(N)+} & \\ \ddots & & \ddots & & & \vdots \\ & \mathbf{H}^{(Q^{(d)}-q)}_{Y-} & \cdots & \cdots & \cdots & \mathbf{H'}^{(Q^{(d)}-q)}_{(q-1)+} \\ & \ddots & \ddots & & & \vdots \\ & & \mathbf{H}^{(Q^{(d)})}_{Y-} & \cdots & \cdots & \mathbf{H'}^{(Q^{(d)})}_{(0)} \end{bmatrix}, \tag{17}$$

where $\mathbf{H}^{(j)}_{i^-}$ represents the net decrement of i packet(s) from the DS buffer when the buffer had j packet(s), $\mathbf{H}^{(j)}_{i^+}$ represents the net increment of i packet(s) from the DS buffer when the buffer had j packet(s). Using eqs. (16) and (17), the inner sub matrices of the queuing model can be defined as follows

$$\mathbf{B}_{L-}^{(E)} = \begin{cases} \mathbf{H}_l^{(e)} &= \alpha_0\beta_{|e-l|}\sum_{\lceil\frac{L+e}{r}\rceil\leq j\leq N_pZ}\mathbf{U}^{(j)}, \text{ if } E{=}L \\ \mathbf{H}_l^{(e)} &= \beta_{|e-l|}\{\sum_{\substack{0\leq j\leq E-L-1 \\ j+L+e \mod r=0}}\alpha_j\mathbf{U}^{(\frac{j+L+e}{r})} \\ & + \alpha_{E-L}\sum_{\lceil\frac{E+e}{r}\rceil\leq j\leq N_pZ}\mathbf{U}^{(j)}\} \end{cases}, \tag{18}$$

$$\mathbf{B}_L^{(0)} = \begin{cases} \mathbf{H}_l^{(0)} &= \beta_l\alpha_L\mathbf{S} \\ \mathbf{H}_0^{(e)} &= \alpha_L\sum_{\substack{0\leq j\leq N \\ j \mod r=0}}\beta_j\mathbf{U}^{(\frac{j}{r})} \\ & + \beta_e\alpha_L\sum_{\lceil\frac{e+1}{r}\rceil\leq j\leq N_pZ}\mathbf{U}^{(j)} \\ \mathbf{H}_{l+}^{(e)} &= \alpha_L\sum_{\substack{l\leq j\leq N \\ j-l \mod r=0}}\beta_j\mathbf{U}^{(\frac{j-l}{r})} \\ \mathbf{H}_q' &= \alpha_L\sum_{q\leq j\leq N}\sum_{\substack{0\leq h\leq j-q \\ h \mod r=0}}\beta_j\mathbf{U}^{(\frac{h}{r})} \\ \mathbf{H}_{l-}^{(e)} &= \begin{cases} \alpha_L\beta_0\sum_{\lceil\frac{l}{r}\rceil\leq j\leq N_pZ}\mathbf{U}^{(j)}, \text{ if } l{=}e \\ \alpha_L\sum_{\substack{0\leq j\leq N \\ j+l \mod r=0}}\beta_j\mathbf{U}^{(\frac{j+l}{r})} \\ +\alpha_L\beta_{e-l}\sum_{\lceil\frac{e+1}{r}\rceil\leq j\leq N_pZ}\mathbf{U}^{(j)} \end{cases} \end{cases}, \tag{19}$$

$$\mathbf{B}_L^{(E)} = \begin{cases} \mathbf{H}_l^{(0)} &= \beta_l \{ \sum_{\substack{0 \le j \le E-1 \\ j \mod r = 0}} \alpha_{j+L} \mathbf{U}^{(\frac{j}{r})} \\ &+ \alpha_{E+L} \sum_{\lceil \frac{E}{r} \rceil \le j \le N_p Z} \mathbf{U}^{(j)} \} \\ \mathbf{H}'_q &= \alpha_L \sum_{q \le j \le N} \sum_{\substack{0 \le h \le j-q \\ h \mod r = 0}} \beta_j \mathbf{U}^{(\frac{h}{r})} \\ \mathbf{H}_{l+}^{(e)} &= \alpha_L \sum_{\substack{l \le j \le N \\ j-l \mod r = 0}} \beta_j \mathbf{U}^{(\frac{j-l}{r})} \\ \mathbf{H}_0^{(e)} &= \beta_e \{ \sum_{\substack{1 \le j \le E \\ j+e \mod r = 0}} \alpha_{j+L} \mathbf{U}^{(\frac{j+e}{r})} \\ &+ \alpha_{E+L} \sum_{\lceil \frac{E+e+1}{r} \rceil \le j \le N_p Z} \mathbf{U}^{(j)} \} \\ &+ \alpha_L \sum_{\substack{0 \le j \le N \\ j \mod r = 0}} \beta_j \mathbf{U}^{(\frac{j}{r})} \\ \mathbf{H}_{l-}^{(e)} &= \begin{cases} \beta_0 \{ \sum_{\substack{0 \le j \le E \\ j+e \mod r = 0}} \alpha_{j+L} \mathbf{U}^{(\frac{j+e}{r})} \\ + \alpha_{E+L} \sum_{\lceil \frac{E+e+1}{r} \rceil \le j \le N_p Z} \mathbf{U}^{(j)} \}, \\ \text{if } l = e \\ \beta_{e-l} \{ \sum_{\substack{1 \le j \le E \\ j+e \mod r = 0}} \alpha_{j+L} \mathbf{U}^{(\frac{j+e}{r})} \\ + \alpha_{E+L} \sum_{\lceil \frac{E+e+1}{r} \rceil \le j \le N_p Z} \mathbf{U}^{(j)} \} \\ + \alpha_L \sum_{\substack{0 \le j \le N \\ j+l \mod r = 0}} \beta_j \mathbf{U}^{(\frac{j+l}{r})} \end{cases} \end{cases}, \tag{20}$$

$$\mathbf{B}_{L-} = \begin{cases} \mathbf{H}_l^{(e)} = \beta_{|e-l|} \sum_{\substack{0 \le j \le M \\ j+L+e \mod r = 0}} \alpha_j \mathbf{U}^{(\frac{j+L+e}{r})}, \end{cases} \tag{21}$$

$$\mathbf{B}_L = \begin{cases} \mathbf{H}_l^{(0)} & = & \beta_l\{ \sum\limits_{\substack{0\le j\le M-L \\ j \mod r=0}} \alpha_{j+L}\mathbf{U}^{(\frac{j}{r})} \\ \mathbf{H}'_q & = & \alpha_L \sum\limits_{q\le j\le N} \sum\limits_{\substack{0\le h\le j-q \\ h \mod r=0}} \beta_j \mathbf{U}^{(\frac{h}{r})} \\ \mathbf{H}_{l+}^{(e)} & = & \alpha_L \sum\limits_{\substack{l\le j\le N \\ j-l \mod r=0}} \beta_j \mathbf{U}^{(\frac{j-l}{r})} \\ \mathbf{H}_0^{(e)} & = & \beta_e \sum\limits_{\substack{L+e\le j\le M \\ j \mod r=0}} \alpha_j \mathbf{U}^{(\frac{j}{r})} \\ & + & \alpha_L \sum\limits_{\substack{0\le j\le N \\ j \mod r=0}} \beta_j \mathbf{U}^{(\frac{j}{r})} \\ \mathbf{H}_{l-}^{(e)} & = & \begin{cases} \beta_0 \sum\limits_{\substack{0\le j\le M-L \\ j+e \mod r=0}} \alpha_{j+L}\mathbf{U}^{(\frac{j+e}{r})}, \\ \text{if } l=e \\ \beta_{e-l} \sum\limits_{\substack{L+1\le j\le M-L-1 \\ j+e \mod r=0}} \alpha_j \mathbf{U}^{(\frac{j+e}{r})} \\ +\alpha_L \sum\limits_{\substack{0\le j\le N \\ j+l \mod r=0}} \beta_j \mathbf{U}^{(\frac{j+l}{r})} \end{cases} \end{cases}, \tag{22}$$

$$\mathbf{J}'_0 = \begin{bmatrix} \mathbf{J_0} & \mathbf{0}_{((Q^{(d)}+1)\times YZ)\times((Q^{(d)}+1)\times X'Z)} \end{bmatrix}, \tag{23}$$

$$\mathbf{J}'_2 = \begin{bmatrix} \mathbf{J_2} \\ \mathbf{0}_{((Q^{(d)}+1)\times X'Z)\times((Q^{(d)}+1)\times YZ)} \end{bmatrix}, \tag{24}$$

$$
\mathbf{B}'_k = \begin{cases}
\mathbf{H}_l^{(0)} &= \beta_d \sum\limits_{k\le j\le M} \sum\limits_{\substack{k\le h\le j\\ h \bmod r=0}} \alpha_j \mathbf{U}^{(\frac{h}{r})} \\
\mathbf{H}_0^{(e)} &= (1-\sum\limits_{0\le j\le k-1}\alpha_j) \sum\limits_{\substack{0\le j\le N\\ j \bmod r=0}} \beta_j \mathbf{U}^{(\frac{j}{r})} \\
&+ \beta_e \sum\limits_{e+k\le j\le M} \sum\limits_{\substack{e\le h\le j-k\\ h+e \bmod r=0}} \alpha_j \mathbf{U}^{(\frac{h+e}{r})} \\
\mathbf{H}'_q &= (1-\sum\limits_{0\le j\le k-1}\alpha_j) \sum\limits_{q\le j\le N} \\
& \sum\limits_{\substack{0\le h\le j-q\\ h \bmod r=0}} \beta_j \mathbf{U}^{(\frac{h}{r})} \\
\mathbf{H}_{l+}^{(e)} &= (1-\sum\limits_{0\le j\le k-1}\alpha_j) \\
& \sum\limits_{\substack{l\le j\le N\\ j-l \bmod r=0}} \beta_j \mathbf{U}^{(\frac{j-l}{r})} \\
\mathbf{H}_{l-}^{(e)} &= \begin{cases}
\beta_l \sum\limits_{k\le j\le M} \sum\limits_{\substack{k\le h\le j\\ h+e \bmod r=0}} \alpha_j \mathbf{U}^{(\frac{h+e}{r})}, \\
\text{if } l{=}e \\
(1-\sum\limits_{0\le j\le k-1}\alpha_j) \\
\sum\limits_{\substack{0\le j\le N\\ j+l \bmod r=0}} \beta_j \mathbf{U}^{(\frac{j+l}{r})} \\
+\beta_{e-l} \sum\limits_{k\le j\le M} \\
\sum\limits_{\substack{k\le h\le j\\ h+e \bmod r=0}} \alpha_j \mathbf{U}^{(\frac{h+e}{r})}
\end{cases}
\end{cases}
, \qquad (25)
$$

$$\mathbf{J}'_1 = \begin{bmatrix} \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & & \\ \mathbf{B}_{1-} & \mathbf{B}_0 & \mathbf{B}_{1+} & \cdots & \mathbf{B}_{M+} & \\ \vdots & & \ddots & & \ddots & \\ \mathbf{B}_{(Y-M+X')-} & \cdots & \cdots & \cdots & & \mathbf{B}'_{(M-1)+} \\ \vdots & & \ddots & & & \vdots \\ \mathbf{B}_{Y-} & \cdots & \cdots & \cdots & \cdots & \mathbf{B}'_{(X'-1)+} \\ \ddots & & \ddots & & & \vdots \\ & & \mathbf{B}_{Y-} & \cdots & \cdots & \mathbf{B}'_{(0)} \end{bmatrix}, , \tag{26}$$

where $0 \leq k \leq M-1$, $0 \leq L \leq M$, $1 \leq L- \leq Y$, $0 \leq l \leq N$, $1 \leq l+ \leq N$, $1 \leq l- \leq Y$, $0 \leq q \leq N-1$, $1 \leq E \leq Q^{(b)}$, $1 \leq e \leq Q^{(d)}$ and $X' = Q^{(b)} - \lfloor \frac{Q^{(b)}}{Y} \rfloor \times Y$.

Appendix C

Let us consider a set of sub matrices $\mathbf{U}^{(x)}(0 \leq x \leq Z)$ and $\mathbf{M}^{(y)}(1 \leq y \leq 4)$. $\mathbf{U}^{(x)}(0 \leq x \leq Z)$ can be written as follows

$$\mathbf{U}_{i,j}^{(x)} = \begin{cases} \mathbf{S}_{i,j} & \text{if} \quad i = x \\ 0 & \text{if} \quad i \neq x \end{cases}, \tag{27}$$

where $0 \leq i, j \leq Z$ and $\mathbf{S}_{i,j}$ denotes the transition probability of transmission rate from state i to state j.

Similarly, $\mathbf{M}^{(y)}(1 \leq y \leq 4)$ can be written as follows

$$\mathbf{M}_{i,j}^{(y)} = \begin{cases} \mathbf{R}_{i,j} & \text{if} \quad i = y \\ 0 & \text{if} \quad i \neq y \end{cases}, \tag{28}$$

where $1 \leq i, j \leq 4$ and $\mathbf{R}_{i,j}$ represents the transition probability of eq. (4.3) from state i to state j. Using eq. (27) and eq. (28) we can define the inner sub matrices of eq. (4.5) as follows

$$\mathbf{H}_0^{(0)} = \beta_0 \mathbf{SR}; \qquad \mathbf{H}_{d+}^{(0)} = \beta_{d+} \mathbf{SR}, \quad 1 \leq d \leq N, \tag{29}$$

$$\mathbf{H}_0^{(e)} = \begin{cases} \beta_0 \mathbf{U}^{(0)}\mathbf{R} + \beta_0 \sum_{\lceil \frac{1}{r} \rceil \le q \le Z} \mathbf{U}^{(q)} \sum_{1 \le y \le 3} \mathbf{M}^{(y)} + \sum_{\substack{1 \le p \le e \\ p \mod r = 0}} \beta_p \mathbf{U}^{(p)}\mathbf{M}^{(4)} \\ + \sum_{\lceil \frac{e+1}{r} \rceil \le x \le Z} \beta_e \mathbf{U}^{(x)}\mathbf{M}^{(4)}, \text{ for } 1 \le e \le N, \\ \beta_0 \mathbf{U}^{(0)}\mathbf{R} + \beta_0 \sum_{\lceil \frac{1}{r} \rceil \le q \le Z} \mathbf{U}^{(q)} \sum_{1 \le y \le 3} \mathbf{M}^{(y)} \\ + \sum_{\substack{1 \le p \le N \\ p \mod r = 0}} \beta_p \mathbf{U}^{(p)}\mathbf{M}^{(4)}, \text{ for } N < e \le Y, \end{cases} \tag{30}$$

$$\mathbf{H}_{d-}^{(e)} = \begin{cases} \beta_0 \sum_{\lceil \frac{e}{r} \rceil \le q \le Z} \mathbf{U}^{(q)}\mathbf{M}^{(4)}, \quad \text{if} \qquad d = e \\ \sum_{\substack{0 \le q \le e-d-1 \\ q+d \mod r = 0}} \beta_q \mathbf{U}^{(q+d)}\mathbf{M}^{(4)} + \beta_{e-d} \sum_{\lceil \frac{e}{r} \rceil \le q \le Z} \mathbf{U}^{(q)}\mathbf{M}^{(4)}, \\ \text{for } e - N \le d \le e - 1, \\ \sum_{\substack{0 \le p \le N \\ p+d \mod r = 0}} \beta_p \mathbf{U}^{(p+d)}\mathbf{M}^{(4)}, \text{for } 1 \le d \le e - N, 1 \le e \le Y - 1, \end{cases} \tag{31}$$

$$\mathbf{H}_{d+} = \beta_d \mathbf{U}^{(0)}\mathbf{R} + \beta_d \sum_{\lceil \frac{1}{r} \rceil \le q \le Z} \mathbf{U}^{(q)} \sum_{1 \le y \le 3} \mathbf{M}^{(y)} + \sum_{\substack{d+1 \le x \le N \\ x-d \mod r = 0}} \beta_x \mathbf{U}^{(x-d)}\mathbf{M}^{(4)}, \ 1 \le d \le N, \tag{32}$$

$$\mathbf{H}_{d-} = \sum_{\substack{0 \le p \le N \\ p+d \mod r = 0}} \beta_p \mathbf{U}^{(p+d)}\mathbf{M}^{(4)}, \ 1 \le d \le Y, \tag{33}$$

$$\mathbf{H}_{d+}^{(e)} = \beta_d \mathbf{U}^{(0)}\mathbf{R} + \beta_d \sum_{\lceil \frac{1}{r} \rceil \le q \le Z} \mathbf{U}^{(q)} \sum_{1 \le y \le 3} \mathbf{M}^{(y)} + \beta_{(e+d)} \sum_{\lceil \frac{e}{r} \rceil \le q \le Z} \mathbf{U}^{(q)}\mathbf{M}^{(4)}$$

$$+ \sum_{\substack{d+1 \leq p \leq d+e-1 \\ p-d \mod r=0}} \beta_p \mathbf{U}^{(p-d)} \mathbf{M}^{(4)},$$

(34)

$$\mathbf{H}_0 \;=\; \beta_0 \mathbf{U}^{(0)} \mathbf{R} + \beta_0 \sum_{\lceil \frac{1}{r} \rceil \leq q \leq Z} \mathbf{U}^{(q)} \sum_{1 \leq y \leq 3} \mathbf{M}^{(y)} + \sum_{\substack{1 \leq p \leq N \\ p \mod r=0}} \beta_p \mathbf{U}^{(p)} \mathbf{M}^{(4)}, \tag{35}$$

$$\begin{aligned} \mathbf{H}'_0 \;&=\; \sum_{0 \leq p \leq N} \beta_p \mathbf{U}^{(0)} \mathbf{R} + \sum_{0`p \leq N} \beta_p \sum_{\lceil \frac{1}{r} \rceil \leq q \leq Z} \mathbf{U}^{(q)} \sum_{1 \leq y \leq 3} \mathbf{M}^{(y)} \\ &+ \sum_{1 \leq j \leq N} \sum_{\substack{1 \leq h \leq j \\ h \mod r=0}} \beta_j \mathbf{U}^{(h)} \mathbf{M}^{(4)}, \end{aligned} \tag{36}$$

$$\begin{aligned} \mathbf{H}'_{k+} \;&=\; \sum_{k \leq p \leq N} \beta_p \mathbf{U}^{(0)} \mathbf{R} + \sum_{k \leq p \leq N} \beta_p \sum_{\lceil \frac{1}{r} \rceil \leq q \leq Z} \mathbf{U}^{(q)} \sum_{1 \leq y \leq 3} \mathbf{M}^{(y)} \\ &+ \sum_{k+1 \leq j \leq N} \sum_{\substack{1 \leq h \leq j-k \\ h \mod r=0}} \beta_j \mathbf{U}^{(h)} \mathbf{M}^{(4)}, \end{aligned} \tag{37}$$

$$\mathbf{F}'_2 = \begin{bmatrix} \mathbf{F_2} \\ \mathbf{0}_{(4 \times X'Z) \times (4 \times YZ)} \end{bmatrix}, \tag{38}$$

$$\mathbf{F}'_0 = \begin{bmatrix} \mathbf{F_0} & \mathbf{0}_{(4 \times YZ) \times (4 \times X'Z)} \end{bmatrix}, \tag{39}$$

$$\mathbf{F}_1' = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{1+} & \cdots & \mathbf{H}_{N+} & & \\ \mathbf{H}_{1-} & \mathbf{H}_0 & \mathbf{H}_{1+} & \cdots & \mathbf{H}_{N+} & \\ \vdots & & \ddots & & \ddots & \\ \mathbf{H}_{(Y-N+X')-} & \cdots & \cdots & \cdots & & \mathbf{H'}_{(N-1)+} \\ \vdots & & \ddots & & & \vdots \\ \mathbf{H}_{Y-} & \cdots & \cdots & \cdots & \cdots & \mathbf{H'}_{(X'-1)+} \\ \ddots & & \ddots & & & \vdots \\ & & \mathbf{H}_{Y-} & \cdots & \cdots & \mathbf{H'}_{(0)} \end{bmatrix}, \tag{40}$$

where

$$X' = Q - \lfloor \frac{Q}{Y} \rfloor \times Y. \tag{41}$$

Appendix D

Using steady state probability vectors of corresponding levels, eq. (2.6) can be written as follows

$$\begin{aligned}
\mathbf{v}_0\mathbf{A_G} + \mathbf{v}_1\mathbf{D_G} &= \mathbf{0} && (42)\\
\mathbf{v}_0\mathbf{C_G} + \mathbf{v}_1\mathbf{E_G} + \mathbf{v}_2\mathbf{F_{2G}} &= \mathbf{0} && (43)\\
\mathbf{v}_1\mathbf{F_{0G}} + \mathbf{v}_2\mathbf{F_{1G}} + \mathbf{v}_2\mathbf{R}_{\text{QBD}}\mathbf{F_{2G}} &= \mathbf{0} && (44)\\
\mathbf{v}_2{\mathbf{R}_{\text{QBD}}}^{(X-4)}\mathbf{F_{0G}} + \mathbf{v}_{X-1}\mathbf{F_{1G}} + \mathbf{v}_X\mathbf{F}'_{\mathbf{2G}} &= \mathbf{0} && (45)\\
\mathbf{v}_{X-1}\mathbf{F}'_{\mathbf{0G}} + \mathbf{v}_X\mathbf{F}'_{\mathbf{1G}} &= \mathbf{0} && (46)
\end{aligned}$$

The vectors of steady state probabilities of repeating levels can be obtained using so-called rate matrix, $\mathbf{R}_{\text{QBD}}$ i.e., $\mathbf{v}_3 = \mathbf{v}_2\mathbf{R}_{\text{QBD}}$, $\mathbf{v}_4 = \mathbf{v}_2\mathbf{R}^2_{\text{QBD}}$,....,$\mathbf{v}_{X-2} = \mathbf{v}_2\mathbf{R}^{(X-4)}_{\text{QBD}}$ [54]. In most application, $\mathbf{R}_{\text{QBD}}$ is computed from an iterative algorithm. One can obtain rate matrix, $\mathbf{R}_{\text{QBD}}$ iteratively as follows

$$\mathbf{F_{0G}} + \mathbf{R}_{\text{QBD}}\mathbf{F_{1G}} + \mathbf{R}^2_{\text{QBD}}\mathbf{F_{2G}} = 0 \qquad (47)$$

Sum of all probabilities should be 1 and can be expressed as follows

$$\mathbf{v}_0\mathbf{1} + \mathbf{v}_1\mathbf{1} + \mathbf{v}_2\mathbf{1} + \mathbf{v}_3\mathbf{1} + ... + \mathbf{v}_{(X-2)}\mathbf{1} + \mathbf{v}_{X-1}\mathbf{1} + \mathbf{v}_X\mathbf{1} = 1 \qquad (48)$$

One can rewrite eq. (48) as follows

$$\mathbf{v}_0\mathbf{1} + \mathbf{v}_1\mathbf{1} + \mathbf{v}_2(\mathbf{I} + \mathbf{R}_{\text{QBD}} + ... + \mathbf{R}^{(X-4)}_{\text{QBD}})\mathbf{1} + \mathbf{v}_{X-1}\mathbf{1} + \mathbf{v}_X\mathbf{1} = 1 \qquad (49)$$

Eq. (49) can be expressed in terms of matrix as follows

$$(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{X-1}, \mathbf{v}_X) \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{II} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix} = 1, \tag{50}$$

where, $\mathbf{II} = (\mathbf{I} + \mathbf{R}_{\text{QBD}} + ... + \mathbf{R}_{\text{QBD}}^{(X-4)})$.

Similarly, eqs. (42)-(46) can be expressed in terms of matrix as follows

$$(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{X-1}, \mathbf{v}_X) \begin{bmatrix} \mathbf{A_G} & \mathbf{C_G} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{D_G} & \mathbf{E_G} & \mathbf{F_{0G}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F_{2G}} & \mathbf{F_{1G}} + \mathbf{R}_{\text{QBD}}\mathbf{F_{2G}} & \mathbf{R}_{\text{QBD}}^{(X-4)}\mathbf{F_{0G}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F_{1G}} & \mathbf{F'_{0G}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F'_{2G}} & \mathbf{F'_{1G}} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{51}$$

Inserting eq. (50) at the end of eq. (51) and eliminate one column that corresponds to the repeating level, eq. (51) can be expressed as follows

$$(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{X-1}, \mathbf{v}_X) \begin{bmatrix} \mathbf{A_G} & \mathbf{C_G} & (\mathbf{0})' & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{D_G} & \mathbf{E_G} & (\mathbf{F_{0G}})' & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{F_{2G}} & (\mathbf{F_{1G}} + \mathbf{R}_{\text{QBD}}\mathbf{F_{2G}})' & \mathbf{R}_{\text{QBD}}^{(X-4)}\mathbf{F_{0G}} & \mathbf{0} & \mathbf{II} \\ \mathbf{0} & \mathbf{0} & (\mathbf{0})' & \mathbf{F_{1G}} & \mathbf{F'_{0G}} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & (\mathbf{0})' & \mathbf{F'_{2G}} & \mathbf{F'_{1G}} & \mathbf{1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{bmatrix} \quad (52)$$

Finally, steady state probability vectors can be obtained from eq. (53).

$$(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{X-1}, \mathbf{v}_X) = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{A_G} & \mathbf{C_G} & (\mathbf{0})^* & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{D_G} & \mathbf{E_G} & (\mathbf{F_{0G}})^* & \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{F_{2G}} & (\mathbf{F_{1G}} + \mathbf{R}_{\text{QBD}}\mathbf{F_{2G}})^* & \mathbf{R}_{\text{QBD}}^{(X-4)}\mathbf{F_{0G}} & \mathbf{0} & \mathbf{II} \\ \mathbf{0} & \mathbf{0} & (\mathbf{0})^* & \mathbf{F_{1G}} & \mathbf{F'_{0G}} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & (\mathbf{0})^* & \mathbf{F'_{2G}} & \mathbf{F'_{1G}} & \mathbf{1} \end{bmatrix}^{-1} \quad (53)$$