2008

Optimal call admission control policies in wireless cellular networks using Semi Markov Decision Process

Wenlong Ni
The University of Toledo

A Dissertation
entitled

Optimal Call Admission Control Policies in Wireless Cellular Networks Using Semi Markov Decision Process

by

Wenlong Ni

As partial fulfillment of the requirements for the
Doctor of Philosophy Degree in Engineering

Advisor: Dr. Mansoor Alam

College of Graduate Studies

The University of Toledo
December 2008

The University of Toledo

College of Engineering

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY

SUPERVISION BY Wenlong Ni

ENTITLED Optimal Call Admission Control Policies in Wireless Cellular Networks Using Semi Markov Decision Process

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

Doctor of Philosophy Degree in Engineering

Dissertation Advisor: Dr. Mansoor Alam

Recommendation concurred by the Committe on Final Examination

Committee: Dr. Biao Ou

Committee: Dr. Ezzatollah Salari

Committee: Dr. Junghwan Kim

Committee: Dr. Mohsin M. Jamali

Committee: Dr. Wei Li(Co-Advisor)

Dean, College of Engineering

An Abstract of

Optimal Call Admission Control Policies in Wireless Cellular Networks Using Semi Markov Decision Process

Wenlong Ni

Submitted in partial fulfillment
of the requirements for the
Doctor of Philosophy Degree in Engineering

The University of Toledo
December 2008

This dissertation investigates the optimal call admission control (CAC) policies for Non Priority Scheme (NPS) and Reserved Channel Scheme (RCS) in wireless cellular networks for the following cases:

1. Homogenous traffic in a cell
2. Multimedia traffic in a cell
3. Homogenous traffic in a network
4. Multimedia traffic in a network

Both new call and handoff call arrival processes are assumed to be Poisson processes, and the call holding times are exponentially distributed with different rate for new calls and handoff calls. Admitting each call would bring a reward to the network provider but holding each call in the system would also incur some cost to the provider. This dissertation focuses on the optimization problem of when to admit or reject a call in order to achieve the maximum reward. By establishing an infinite horizon discounted Semi Markov Decision Process (SMDP) model, this research verifies that the optimal policies of all the four cases mentioned above are state-related control

policies for NPS as well as RCS. More over, this research also finds that optimal control policy for Case (1) above is still state-related control policy even if the combined arrival process for new and handoff calls are described by General Distribution.

The numerical results presented in this dissertation in both tables and diagrams for all of the above cases are consistent with the theoretical results derived.

Acknowledgments

This is a great opportunity to express my respect to all those people who have made this dissertation possible and because of them my graduate experience has been one that I will cherish forever.

My deepest gratitude is to my advisor Dr. Mansoor Alam. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on myself and at the same time the guidance to recover when my steps faltered. I am also thankful to him for encouraging the use of correct grammar and consistent notation in my writings and for carefully reading and commenting on revisions of this manuscript. I hope that one day I would become as good an advisor to my students as Alam has been to me.

I truthfully want to thank my co-advisor Dr. Wei Li for inspiring me to this area of research. Dr. Li has been always there to listen and give advice. I am deeply grateful to him for the long discussions that helped me sort out the technical details of my work. Li taught me how to question thoughts and express ideas. His patience and support helped me overcome many crisis situations and finish this dissertation. Dr. Li's insightful comments and constructive criticisms at different stages of my research were thought-provoking and they helped me focus my ideas. I am grateful to him for holding me to a high research standard and enforcing strict validations for each research result, and thus teaching me how to do research.

I would like to acknowledge Dr. Biao Ou, Dr. Ezzatollah Salari, Dr. Junghwan Kim, and Dr. Mohsin M. Jamali for necessary assistance like discussions on related topics and their teachings that helped me improve my knowledge in the area.

Most importantly, none of this would have been possible without the love and patience of my family. This dissertation is dedicated to them. I would like to express my heart-felt gratitude to my parents Shaocheng/Qingyu(Wen), my sisters Jian/Jun/Wei and especially to my wife Wuming(Xu). It is their understandings and supports that helped me overcome setbacks and be focused on my graduate study through these difficult years.

Finally, I appreciate the financial support from NSF that funded parts of the research discussed in this dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Literature Survey

A cellular network is a radio network made up of a number of radio cells (or just cells) each served by a fixed transmitter, known as a cell site or base station. These cells are used to cover different areas in order to provide radio coverage over a wider area than the area of one cell. Cellular networks are inherently asymmetric with a set of fixed main transceivers each serving a cell and a set of distributed (generally, but not always, mobile) transceivers which provide services to the network's users.

The tremendous growth of the wireless/mobile user population, coupled with the bandwidth requirements of multimedia applications, requires efficient reuse of the scarce radio spectrum allocated to wireless/mobile communications. Efficient use of bandwidth is also important from a cost of service point of view. Cellular networks offer a number of advantages over alternative solutions:

- increased capacity
- reduced power usage
- better coverage

As shown in Figure (1-1), the mobile users in a cell are served by a base station. Before a mobile user can communicate with other user(s) in the network, a connection must usually be established between the users. The establishment and maintenance of a connection in a wireless network is the responsibility of the base station. To establish a connection, a mobile user must first specify its traffic characteristics and quality of service (QoS) needs. For example, in a cellular phone network, the traffic characteristics and the QoS needs of voice connections are known a priori to the base station, and therefore, they are usually implicit in a connection request.

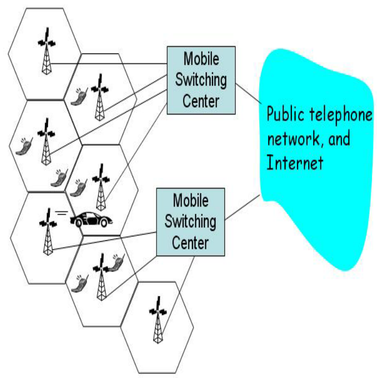

Figure 1-1: Wireless Cellular Network

The advent of the third generation of wireless multimedia services, brought about the need to adapt the existing mobile cellular networks to make them carry the different types of services. The wireless device could be a wireless telephone, personal digital assistant (PDA), Palm Pilot, laptop with wireless card, or Web-enabled phone. For simplicity, they all can be called Mobile Station (MS).

There have been extensive studies on wireless cellular networks, [1,7–11,17]. The wireless cellular network has to be able to support multiple classes of traffic with different Quality of Service (QoS) requirements, i.e., different number of channels needed, holding time of the connection and cell residence time. When there is a call coming into a wireless network, the system must decide whether to admit the call or block it. What makes the decision difficult is that, to achieve some sense of optimality, one needs to consider the future status of the network resources and the pattern of the future arrival requests. For example, even if the system currently has sufficient resources to handle the call, admitting the call may result in other, already in-progress calls being dropped in the future. It becomes an optimization control problem of when to reject or admit the calls to achieve the maximum reward. As a result, much research has been devoted to call admission control, [3,4,14,22]. Markov Decision Process (MDP) is a well known technique for this type of research.

1.1 Call Admission Control Policy

Call admission control (CAC) is the practice or process of regulating traffic volume in communications networks, particularly in wireless mobile networks and in Internet telephony. CAC is a concept that applies only to real time media traffic and not to data traffic. CAC mechanisms complements the capabilities of QoS tools to protect traffic from the negative effects of other traffic and to keep excess traffic off the network. Call Admission Control(CAC) is used to prevent congestion in traffic. It is a preventive Congestion Control Procedure. It is used in the Call Setup phase. CAC can also be used to ensure, or maintain, a certain level of quality in communications networks, or a certain level of performance in Internet nodes and servers.

Most CAC algorithms work by regulating the total utilized bandwidth, the total number of calls, or the total number of packets or data bits passing a specific point per unit time. If a defined limit is reached or exceeded, a new call may be prohibited from entering the network until at least one current call terminates. Alternatively, a graceful degradation methodology can be implemented. This means that the audio quality of individual calls can deteriorate to a certain extent before new calls are denied entry. Another method involves the regulation of calls according to defined characteristics such as priority descriptors. Still another method prevents new calls from entering the network only if the resources of the server would be overburdened by such calls, as shown in Figure (1-2).

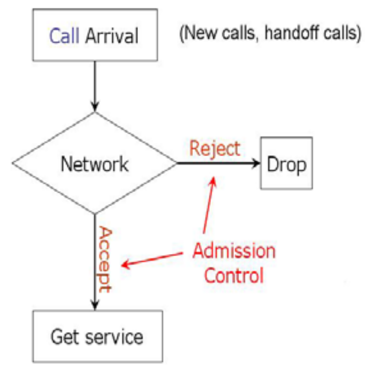

Figure 1-2: Call Admission Control

In a wireless cellular network, when a user moves from one cell to another, the base station in the new cell must take responsibility for all the previously established connections. A significant part of this responsibility involves allocating sufficient resources in the cell to maintain the needs of the established connection(s). If sufficient resources are not allocated, which may result in premature termination of the connection. Since premature termination of established connections is usually more objectionable than rejection of a new connection request, it is widely believed that a wireless cellular network must give higher priority to the handoff connection requests as compared to new connection requests. Many different admission control strategies have been discussed in literature to provide priorities to handoff requests without significantly jeopardizing the new connection requests.

It is well known that the call admission policy have a big impact on the performance of a wireless cellular network. In this dissertation, call admission control problem in a cellular network is modeled as a semi-Markov decision process (SMDP), [13, 16, 25, 32], which models the system evolution in continuous time. In SMDP, action choice determines the joint probability distribution of the subsequent state and the time between decision epochs. To achieve the optimality on the rewards, the SMDP model is described as an infinite-horizon problem with discounting. By using the Rate Uniformization technique [13], the continuous time SMDP model can be transformed to a discrete MDP model, thus the theorems and algorithms for MDP models could be applied. With a finite state space and a finite action space, the Value Iteration Method is used to solve the SMDP problem, not in the way of Linear Programming as in [4]. [7] proposed a similar model but it only studied on NPS and no holding costs. More information of "Semi Makkov Decision Process" are described in the chapter 2.

1.2 Resource Allocation Schemes

Channels in a wireless communication system typically consist of time slots, frequency bands and/or CDMA pseudo noise sequences, but in an abstract sense, they can represent any generic transmission resource. Channel allocation deals with the allocation of channels to cells in a cellular network. Cells allow users to communicate via the available channels to a specific channel assignment scheme.

In celluar networks, the calls can be divided into two groups: new calls and handoff calls (calls moving from one cell to another). There are several channel allocation schemes of allocating a channel upon a new call arrival or handoff attempt, like Non Priority Scheme, Reserved Channel Scheme (RCS), Queueing Priority Scheme, etc. Much work has been done on these schemes to effectively utilize the resources in the wireless networks, [23, 24]. The advantages and disadvantages of these schemes are shown in the below.

1. Non-Prioritized Scheme (NPS): This scheme does not differentiate handoff calls from new calls. If there are channels available for a request, one channel is assigned to the request. If there are no free channels, the request is rejected immediately and we have a forced termination or a blocked call. Remark: It is applicable if the forced termination of handoff calls is not important, and the implementation cost is low.

2. Reserved Channel Scheme (RCS): This scheme divides channels into two groups. One group can be assigned to both the initial access and the handoff access. The other group can only be assigned to the handoff access. The purpose is to reduce the probability of the forced termination. Remark: The new call-blocking probability for RCS is large than that of NPS, RCS is easy to implement, and it reduces the forced terminal probability more effectively than NPS, and the implementation cost is not high.

3. Queuing Priority Scheme (QPS): Because adjacent cells are overlapped, the Mobile Station (MS) in the overlapped area can use channels in either of the adjacent cells. The overlapped area is called the handoff area. When an MS is in the handoff area, if the destination cell has no free channels, the MS maintains the existing channel of the source cell. The handoff request is queued and sent to the base station of the destination cell. If a channel in the destination cell is available before the MS crosses the handoff area, the channel is assigned to the MS. Otherwise, the call is forced terminated. Remark: QPS effectively reduce forced terminations, at the expense of increased new call blocking, the probability of incomplete call from QPS is slightly lower than that for NPS. It is good when the BS density is high.

4. Sub-Rating Scheme (SRS): When an MS is in the handoff area and the destination cell has no free channels, an occupied channel in the destination cell is divided into two sub-channels, one for the original call and one for the handoff call. These two sub-channels will become two channels again when there are channels available. Remark: Compared with other schemes, SRS has the least forced termination probability. The benefit is gained at the expense of extra hardware/software complexity required to sub-rate a channel. SRS has the least probability of incomplete calls. It is best if the voice degradation does

not matter much.

A good resource allocation algorithm is the one that yields high spectral efficiency for a specified grade of service (including link quality, probability of new call blocking, and the probability of forced termination) and given degree of computational complexity. In this dissertation we consider a lot on RCS schemes and some about NPS schemes.

1.3 Contributions

In this dissertation, we consider a wireless network supporting diverse traffic characteristics of voice, data, and video applications under certain assumptions of calls arrival process and the probability distribution of calls service time. Accepting each call would contribute a reward to the system and the system incurs a holding cost per unit time for the calls. This dissertation builds several models of wireless cellular networks. In the first three models we consider a single cell with homogeneous traffic. In the first model we take all the incoming calls (new calls, handoff calls) to a cell as one class, the calls arrival process follows a General distribution. In the following models we treat new calls and handoff calls separately, the calls arrival processes are Poisson processes.

We continue our models to multimedia traffic and wireless network with multiple cells. For multimedia traffic, we consider there are many types of calls requiring with different Quality of Service (QoS) requirements, i.e., different number of channels needed, holding time of the connection and cell residence time. The call connections can differ in the amount of resources (say bandwidth) required to meet their needs. For a wireless network with multiple cells, we consider that there are routing probabilities between cells and the way how we deal with the handoff calls inside the network.

The question is when to accept/reject the different incoming calls to achieve the maximum reward or to derive the optimal call admission control policy. By using SMDP technique to solve the optimization problem, we not only figure out the optimal call admission control policy for the maximum reward, but also verify that under certain conditions of cost functions, the optimal call admission control policy is a control limit or threshold policy. Theoretical results are made on the optimal call admission control policies based on the property of cost functions. Numerical analysis and simulation is listed and discussed after the model formulation which demonstrates the applicability of the theoretical results. The proposed approach and the results derived are applicable in real life for the design and performance analysis of wireless cellular networks.

Chapter 2

Background

This chapter provides some necessary background information which are related to this dissertation.

2.1 Wireless Cellular Network

The most common example of a cellular network is a mobile phone (cell phone) network. A mobile phone is a portable telephone which receives or makes calls through a cell site (base station), or transmitting tower. Radio waves are used to transfer signals to and from the cell phone. Large geographic areas (representing the coverage range of a service provider) are split up into smaller cells to deal with line-of-sight signal loss and the large number of active phones in an area. In cities, each cell site has a range of up to approximately 0.5 mile, while in rural areas, the range is approximately 5 miles. Many times in clear open areas, a user may receive signal from a cell 25 miles away. Each cell overlaps other cell sites. All of the cell sites are connected to cellular telephone exchanges "switches", which in turn connect to the public telephone network or another switch of the cellular company.

Compared to other mobile networks, wireless cellular networks have two features as shown below:

1. Handoff: The use of multiple cells means that, if the distributed transceivers are mobile and moving from place to place, they also have to change from cell to cell. The mechanism for this differs depending on the type of network and the circumstances of the change. As the phone user moves from one cell area to another, the switch automatically commands the handset and a cell site with a stronger signal (reported by the handset) to go to a new radio channel (frequency). When the handset responds through the new cell site, the exchange switches the connection to the new cell site.

2. Frequency Reuse: The increased capacity in a cellular network, compared with a network with a single transmitter, comes from the fact that the same radio frequency can be reused in a different area for a completely different transmission. If there is a single plain transmitter, only one transmission can be used on any given frequency. Unfortunately, there is inevitably some level of interference from the signal from the other cells which use the same frequency. This means that, in a standard FDMA system, there must be at least a one cell gap between cells which reuse the same frequency. Modern mobile phones use cells because radio frequencies are a limited, shared resource. Cell-sites and handsets change frequency under computer control and use low power transmitters so that a limited number of radio frequencies can be reused by many callers with less interference.

Currently the biggest challenges on wireless cellular networks are mainly due to the limited resources while providing diverse different kinds of services. One remarkable challenge is how to handle handoff calls which result in the call admission control, which is the focus of study in this dissertation.

2.2 Markov Decision Process

Markov decision processes (MDPs) provide a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of the decision maker. MDPs are useful for studying a wide range of optimization problems solved via dynamic programming and reinforcement learning. MDPs were known at least as early as the 1950s (cf. Bellman 1957). Today they are used in a variety of areas, including robotics, automated control, economics and in manufacturing.

More precisely a Markov Decision Process is a discrete time stochastic control process characterized by a set of states; in each state there are several actions from which the decision maker must choose. For a state s and an action a, a state transition function determines the transition probabilities to the next state. The decision maker earns a reward for each state visited. The states of an MDP possess the **Markov property**: If the current state of the MDP is known, transitions to a new state are independent of all previous states.

A Markov decision process model consists of five elements:

- decision epochs
- states
- actions
- transition probabilities
- rewards.

Decisions are made at time points referred to as *decision epochs*. At each decision epoch, the system occupies a *state*. We denote the set of possible system states by S. If, at some decision epoch, the decision maker observes the system in state $s \in S$, he may choose action a from the set of allowable actions in s, A_s. As a result of choosing action $a \in A_s$ in state s at decision epoch t, the decision maker receives a reward and the system state at the next decision epoch is determined by the transition probability distribution.

A *decision rule* prescribes a procedure for action selection in each state at a specified decision epoch. A *policy* or *strategy* (in this dissertation we would always say policy) specifies the decision rule to be used at all decision epochs. The solution to a Markov Decision Process can be expressed as a policy π, which gives the action to take for a given state, regardless of prior history.

In this dissertation, the Semi Markov decision process (SMDP) model is used for system modeling, which generalizes the MDP in the following characteristics

- SMDP allows the decision maker to choose actions whenever the system state changes.
- SMDP models the system evolution in continuous time.
- SMDP allows the time spent in a particular state to follow an arbitrary probability distribution.

In SMDP, action choice determines the joint probability distribution of the subsequent state and the time between decision epochs. So the MDP can be viewed as a special case of SMDP.

In SMDP models, decision rules range in generality from deterministic Markovian to randomized History Dependent, depending on how they incorporate past information and how they select actions. Deterministic Markovian decision rules are functions $d_t : S \to A_s$, which specify the action choice when the system occupies state s at decision epoch t. A policy is stationary if, for each decision epoch t, $d_t = d$ is the same, which can be denoted by d^∞. For each policy π, let $v_\alpha^\pi(s)$ denote the total expected infinite-horizon discounted reward with α as the discount factor, given that the process occupies state s at the first decision epoch. Our objective is to find an optimal policy π that can bring the maximum total expected discounted reward $v_\alpha^\pi(s)$ for every initial state s.

2.3 Poisson Process

In teletraffic theory, often the arrival process of customers can be described by a Poisson process, the customers may be calls or packets. Poisson process is a viable model when the calls or packets originate from a large population of independent users.

A Poisson process is a stochastic process which is used for modeling random events in time that occur to a large extent independently of one another (the word event used here is not an instance of the concept of event frequently used in probability theory).

In its most general form, the only two conditions for a stochastic process to be a Poisson process are:

1. **Memorylessness:** The numbers of changes in nonoverlapping intervals are independent for all intervals.

2. **Orderliness:** The probability of two or more changes in a sufficiently small interval h is essentially 0.

Poisson process is one of the most important models used in queueing theory. It can be defined in three different (but equivalent) ways:

1. Poisson process is a pure birth process: In an infinitesimal time interval dt there may occur only one arrival. This happens with the probability λdt independent of arrivals outside the interval.

2. The number of arrivals N(t) in a finite interval of length t obeys the Poisson(λdt) distribution,

$$PN(t) = n = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

, Moreover, the number of arrivals N(t_1, t_2) and N(t_3, t_4) in non-overlapping intervals ([t_1, t_2], [t_3, t_4]) are independent.

3. The interarrival times are independent and obey the Exponential distribution: P$\{interarrivaltime > t\} = e^{-\lambda t}$

2.4 Channel Modeling

How to model a wireless cellular network? The calls are coming and leaving at random time points making the call arrival process to be a random process. In this section we will show you a generic system model on a wireless cellular network.

If we make assumptions like below:

1. All MSs are assumed to be uniformly distributed through the cell.

2. Each MS moves at a random speed and to an arbitrary random direction.

Then we can assume the calls arrival process is a Poisson process and again making the following assumptions.

1. The arrival rate of originating call is given by λ_N.

2. The arrival rate of handoff call is given by λ_H.

3. The call service rate is given by μ.

With this system model we can calculate the blocking probabilities for new calls and handoff calls which gives us a way to design the cellular network to meet the Qos demands. We can see that this is a very simple model. In this dissertation we

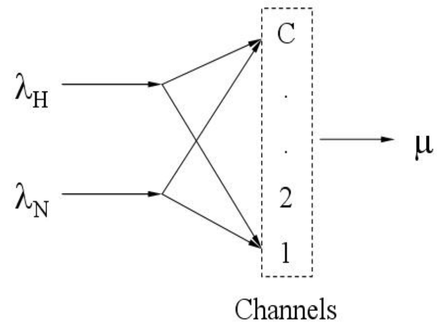

Figure 2-1: A generic System Model

made much more complex models, we not only studied the different arrival processes other than the Poisson processes, but also studied the performance of a multimedia network with multiple cells.

Chapter 3

Modeling, Simulation and Analysis of Homogeneous Traffic In a Cell

Cellular networks are defined by static base stations which divide the fields into cells. All radio communication is between these base stations and the clients. Usually, each static base station forwards and receives packets to other base stations by another (hard-wired) network. Regarding movement of clients one is only interested in whether the node enters or leaves a cell. It is not interesting where the node is exactly located within a cell. Usually network cells overlap and so, there are areas where a client can choose among several base stations. One of the main mobility problems and applications for cellular networks is Cellular Handoff: To provide a robust protocol that allows to move between cells without interrupting and disturbing communication.

3.1 Introduction

In this chapter we study the optimal call admission control policy on a single cell in a wireless cellular network. First, we study on the case as if there is only a single type of traffic (voice, video, etc) in the network. Each call coming into the cell requires a certain range of bandwidth which we call it as a channel and the cell has a number of channels to serve. As we have shown in previous chapters, in wireless cellular networks, the calls coming to a cell can be divided into two groups: new calls and handoff calls. Many research studies devoted to call admission control [3, 4, 13] assume that the arrival processes of new calls and handoff calls are Poisson processes. But by the trend in decreasing cell sizes, the number of handoffs experienced during a typical connection are increased, hence increasing the complexity of dependence between new calls and handoff calls. This makes the modeling of wireless networks more difficult. In this chapter, we build several SMDP models on a single cell with one type of calls. In the first model, we take all the incoming calls to a single cell as if the calls all belong to one class. The interarrival process of calls follows a General distribution. In the following models, we assume the arrival processes of new calls and handoff calls are Poisson processes. And we consider both the RCS schemes and the NPS schemes. Compared to other resource allocation schemes, the RCS reserves some channels for handoff calls, which reduced the probability of the forced termination. While in the NPS Schemes, new calls and handoff calls are treated equally. Any incoming call can get a service if there are free channels.

In each model we assume there is some reward for accepting each call and there is holding costs for calls staying in the system per unit time. As we have learned through the numerical analysis and theoretical deductions, the optimal CAC policies are control limit policies if the cost functions have some special properties.

We made three models to study the performance of homogeneous traffic on a cell in the cellular network. In the first model we took all the calls coming into a cell as one class. In the second and third models we separate the calls as the new calls and handoff calls, then we studied their performances under NPS and RCS schemes.

3.2 One class model

In this section, we do not differentiate between new calls and handoff calls. We take all the incoming calls to a single cell as if the calls all belong to one class. We do not assume the exact format of calls arrival processes.

3.2.1 Model Formulation

To build the model, other assumptions are shown in the below:

1. The arrival instants of calls follow an interarrival distribution A(t).

2. There are m channels and C-m buffer ($C \geq m$)in the system. The service in each channel is independent. The service time of each call follows a negative exponential distribution with rate μ.

3. Accepting each incoming call would contribute R units of reward to the system. The system incurs a cost at rate $f(j)$ per unit time when there are j calls in the system, $f(j) = 0$ if $j \leq m$.

Based on the above assumptions, considering the five elements of a SMDP model, we can build the SMDP model for the system as follows:

1. Let the state variable denote the number of calls in the system at any time point, so the state space is $S = \{0, 1, \ldots, C\}$.

2. Let action a_R denote rejecting a call arrival, action a_A correspond to accepting a call arrival. So for each state s, the action set is $A_s = \{a_R, a_A\}$.

3. The decision epochs are those time points when an arrival comes to the system. Let $F(t|s,a)$ denote the probability the next decision epoch occurs within t time units when the system is in state s and taking action a. We have $F(t|s,a) = A(t)$.

4. As a consequence of choosing action $a \in A_s$, let the next decision epoch occur at or before time t, and the system state at the next decision epoch equal j with probability $Q(t, j|s, a)$, and $Q(dt, j|s, a)$ for the time differential. Also, let $p(j|t, s, a)$ denote the probability that the process occupies state j, t time units after a decision epoch in state s, action a was chosen and that the next decision epoch has not occurred prior to time t. From *Chp XI.3* in [8], for action $a = a_R$, we have

$$p(j|t, s, a_R) = \begin{cases} e^{-m\mu t}\frac{(m\mu t)^{s-j}}{(s-j)!}, & s \geq j \geq m, \\ b_{s-j}(s, 1 - e^{-\mu t}), & j \leq s \leq m, \\ \int_0^t b_{m-j}(m, 1 - e^{-\mu(t-y)})E_i(dy), & s > m > j, \end{cases}$$

where m is the number of channels, μ is the service rate of each channel, $i = s + 1 - m$. E_k denotes the *Erlang distribution with k stages and intensity μm*, and $b_k(n, p)$ is the binomial probability of $\begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{n-k}$. Next, let

$$p(0|t, s, a_R) = 1 - \sum_{j \neq 0, j \leq s} p(j|t, s, a_R).$$

Since admitting a job immediately migrates the system from state s to state $s + 1$, we have $p(j|t, s, a_A) = p(j|t, s + 1, a_R)$.

By referring to the *embedded Markov Decision process*, which describes the evolution of system at decision epochs only. Let $P(j|s, a)$ denote the probability that the embedded Markov Decision process occupies state j at the subsequent decision epoch when action a is chosen in state s at the current decision epoch. It follows that

$$P(j|s, a) = \int_0^\infty p(j|t, s, a)dA(t).$$

And, from [13], we have

$$Q(t,j|s,a) = P(j|s,a)F(t|s,a) = P(j|s,a)A(t).$$

5. For $s \in S$ and $a \in A_s$, let $r(s,a)$ denote the expected total discounted reward between two decision epochs, given the system is in state s and action a is chosen. Then, we have

$$r(s,a) = k(s,a) + \int_0^\infty \sum_{j\in S} \left[\int_0^t e^{-\alpha x} c(j,s,a)p(j|x,s,a)dx\right] dA(t), \tag{3.1}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_R, \\ R, & a = a_A, \end{cases}$$

and

$$c(j,s,a) = \begin{cases} -f(j), & a = a_R, \\ -f(j+1), & a = a_A. \end{cases}$$

Similarly, since admitting a job immediately migrates the system from state s to state $s+1$, we have $r(s,a_A) = R + r(s+1,a_R)$.

For the admission control problem, both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy. And for a stationary policy π, we have

$$v_\alpha^\pi(s) = r(s,d(s)) + \sum_{j\in S} \int_0^\infty e^{-\alpha t} Q(dt,j|s,d(s)) v_\alpha^\pi(j).$$

In general, define

$$m(j|s,a) \equiv \int_0^\infty e^{-\alpha t} Q(dt,j|s,a) \equiv P(j|s,a)\int_0^\infty e^{-\alpha t} dA(t),$$

for $j \in S$, $s \in S$ and $a \in A_s$.

Since $m(j|s, a_A) = m(j|s+1, a_R)$, the optimality equations that achieve the maximum reward with initial state s become

$$v(s) = \max\left\{r(s, a_R) + \sum_{j=0}^{s} m(j|s, a_R)v(j),\ \ R + r(s+1, a_R) + \sum_{j=0}^{s+1} m(j|s+1, a_R)v(j)\right\}. \tag{3.2}$$

3.2.2 Theoretical Results

Further, we show that the optimal policy is a control limit policy or a threshold policy. The control limit policy means that its decision rule is defined as

$$d(s) = \begin{cases} a_A, & s \le K, \\ a_R, & s > K, \end{cases}$$

where $K \ge 0$ is a constant. From equation (3.2), the value of $v(s)$ takes the maximal value from two parts. Define

$$v(s, a_R) = r(s, a_R) + \sum_{j=0}^{s} m(j|s, a_R)v(j),$$

so equation (3.2) can be written as

$$v(s) = \max\left[v(s, a_R), R + v(s+1, a_R)\right]. \tag{3.3}$$

Theorem 3.1 *If the function $v(s, a_R)$ is concave and nonincreasing, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let d be its decision rule, $\triangle v(s, a_R) = v(s+1, a_R) - v(s, a_R)$, so $\triangle v(s, a_R) \le 0$

and is nonincreasing. From equation (3.2), we have

$$d(s) = \begin{cases} a_A, & \triangle v(s, a_R) > -R, \\ a_R, & \triangle v(s, a_R) \leq -R. \end{cases}$$

So, if $d(s) = a_R$, we have $d(s+1) = a_R$, and so on. Consequently the optimal policy is a control limit policy (or threshold policy).

Next, let $c = \int_0^\infty e^{-\alpha t} dA(t)$, we have

$$\sum_{j \in S} m(j|s, a) = c.$$

Theorem 3.2 *$v(s)$ has the same value for states $s \leq m-1$ if there is only waiting cost in the system.*

Proof: By definition for the waiting cost, from equation (3.1) we have $r(s, a_R) = 0$, $s \leq m$. Using Value Iteration Method, let $v^0(s) = 0$. From equation (3.2) and (3.3) we have $v^1(s) = R$, $v^2(s) = R + c * R$. As the iteration continues, we have $v^n(s) = R + c * R + \cdots + c^{n-1} * R$. As n goes to ∞, we get

$$v(s) = \lim_{n \to \infty} v^n(s) = \frac{R}{1-c},$$

for states $s \leq m-1$.

Next, to make the equations simpler, define

$$q_k = \int_0^\infty e^{-m\mu t} \frac{(m\mu t)^k}{k!} dA(t).$$

So for action $a = a_R$, $s \geq j \geq m$, we have $P(j|s, a_R) = q_{s-j}$ and $m(j|s, a_R) = c * q_{s-j}$. Then for states $s \geq m$, we have

$$v(s, a_R) = r(s, a_R) + c * \left[\sum_{j=m}^{s} q_{s-j} v(j) + (1 - \sum_{j=m}^{s} q_{s-j}) v(m-1)) \right]. \tag{3.4}$$

Next, define

$$p_k = \int_0^\infty \int_0^t e^{-\alpha x} e^{-m\mu x} \frac{(m\mu x)^k}{k!} dx dA(t).$$

So, we have

$$\sum_{k=0}^{\infty} p_k = \frac{1}{\alpha}(1-c),$$

which is also a constant. For action $a = a_R$ and states $s \geq m$, equation (3.1) can be rewritten as

$$r(s, a_R) = -\sum_{j=m}^{s} f(j) p_{s-j}. \tag{3.5}$$

Theorem 3.3 *If the cost function $f(j)$ is convex and nondecreasing on j, the optimal policy is a control limit policy.*

Proof: First, from equation (3.5), for states $s \geq m$ we have

$$\triangle r(s, a_R) = r(s+1, a_R) - r(s, a_R) = -\sum_{j=m}^{s} \triangle f(j) p_{s-j},$$

which is nonincreasing. By definition $r(s, a_R)$ is concave nonincreasing.

We use Value Iteration Method to show that for states $s \geq m$, $v(s, a_R)$ is also concave and nonincreasing.

1. Set $v^0(s) = 0$ and substitute this into equation (3.2) to obtain $v^1(s, a_R) = r(s, a_R)$. From equation (3.3), by lemma A.1 , $v^1(s)$ is concave nonincreasing.

2. Set n=1, according to equation (3.4),

$$\triangle v^{n+1}(s, a_R) = \triangle r(s, a_R) + c * \left[\sum_{j=m}^{s} \triangle v^n(j) q_{s-j} \right].$$

From lemma A.3, $v^{n+1}(s, a_R)$ is concave nonincreasing. Next, from lemma A.1, $v^{n+1}(s)$ is concave nonincreasing.

3. Set n=n+1, then go back to step 2.

4. As the iteration continues, as n goes to ∞, $v^n(s, a_R)$ is always concave nonincreasing, so is $v^n(s)$. By the *Theorem 11.3.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^n(s)$ converges to $v(s)$, thus $v(s)$ is concave nonincreasing.

5. The verification is over. From *Theorem 3.1* the optimal policy is a control limit policy.

3.2.3 Simulation Results

The simulation parameters are shown in Table 3.1.

Table 3.1: Parameters setting

λ	μ	R	C	m	α
2	2	1	10	2	0.1

Next, set the waiting cost function $f(j)$ as

$$f(j) = \begin{cases} 0.2 * (j - m), & j > m, \\ 0, & j \le m, \end{cases}$$

which is convex nondecreasing and fits the theorem requirement. We studied two cases of different interarrival processes $A(t)$. In the first case we suppose interarrival process is a Poisson process, in the second case it is an $Erlang-k$ distribution.

•Case I: Poisson Process

Using the Value Iteration Method, based on the simulation parameters and cost function, we got the values of $V(s,a)$ as shown in Table 3.2.

Table 3.2: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	1.7521	1.7521	1.6281	1.0703	0.2177	-0.8255
$a = a_A$	2.6281	2.6281	2.0703	1.2177	0.1745	-1.0206
$v(s)$	2.6281	2.6281	2.0703	1.2177	0.2177	-0.8255
v(s,a)	6	7	8	9	10	
$a = a_R$	-2.0206	-3.3394	-4.7605	-6.2671	-7.3144	
$a = a_A$	-2.3394	-3.7605	-5.2671	-7.2671	-8.4545	
$v(s)$	-2.0206	-3.3394	-4.7605	-6.2671	-7.3144	

Next, the value of μ is changed to $\mu = 0.4$, which makes the system in a heavy traffic load, and we see what happens to the values of $V(s,a)$ as shown in Table 3.3.

Table 3.3: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	1.9269	1.9269	1.8903	1.6161	1.1424	0.4994
$a = a_A$	2.8903	2.8903	2.6161	2.1424	1.4994	0.7116
$v(s)$	2.8903	2.8903	2.6161	2.1424	1.4994	0.7116
v(s,a)	6	7	8	9	10	
$a = a_R$	-0.2884	-1.1988	-2.2045	-3.2949	-4.4612	
$a = a_A$	-0.1988	-1.2045	-2.2949	-4.2949	-4.6947	
$v(s)$	-0.1988	-1.1988	-2.2045	-3.2949	-4.4612	

Next, the value of α is changed to $\alpha = 0.2$, and the values of $V(s,a)$ as shown in Table 3.4.

From these tables we see that both $v(s,a_R)$ and $v(s,a_A)$ are concave nonincreasing on s, so the deterministic stationary decision rule of optimal policy is and the optimal policy is a control limit policy. All these results testify our theorems.

Table 3.4: Total Expected Discounted Reward $V(s, a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	0.9348	0.9348	0.8695	0.478	-0.1099	-0.8215
$a = a_A$	1.8695	1.8695	1.478	0.8901	0.1785	-0.6145
$v(s)$	1.8695	1.8695	1.478	0.8901	0.1785	-0.6145
v(s,a)	6	7	8	9	10	
$a = a_R$	-1.6145	-2.4633	-3.3531	-4.2809	-5.1036	
$a = a_A$	-1.4633	-2.3531	-3.2809	-5.2809	-4.502	
$v(s)$	-1.4633	-2.3531	-3.2809	-4.2809	-4.502	

Next, if we the cost function as $f() = 1$ constantly, the values of $V(s, a)$ as shown in Table 3.5.

Table 3.5: Total Expected Discounted Reward $V(s, a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	0	0	-1.8367	-3.2362	-4.273	-5.0535
$a = a_A$	-0.8367	-0.8367	-2.2362	-3.273	-4.0535	-4.6553
$v(s)$	0	0	-1.8367	-3.2362	-4.0535	-4.6553
v(s,a)	6	7	8	9	10	
$a = a_R$	-5.6553	-6.1361	-6.5593	-7.059	-6.7788	
$a = a_A$	-5.1361	-5.5593	-6.059	-8.059	-4.3117	
$v(s)$	-5.1361	-5.5593	-6.059	-7.059	-4.3117	

As seen from Table 3.5, there is no limit in the decision rule, so the optimal policy is not a control limit policy.

•**Case II A(t):** $Erlang - k$ **distribution**

Suppose interarrival process $A(t)$ is an $Erlang - k$ distribution with $k = 5$, which is a good estimate of the real process. The probability density function $dA(t)$ then become $dA(t) = \frac{\lambda^5 t^4 e^{-\lambda t}}{4!}$

Based on the simulation parameters and cost function, we got the values of $V(s, a)$ as shown in Table 3.6.

Next, the value of μ is changed to $\mu = 0.4$, which makes the system in a heavy traffic load, and we see what happens to the values of $V(s, a)$ as shown in Table 3.7.

Table 3.6: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	1.9978	1.9978	1.9967	1.5917	0.8584	-0.1449
$a = a_A$	2.9967	2.9967	2.5917	1.8584	0.8551	-0.3735
$v(s)$	2.9967	2.9967	2.5917	1.8584	0.8584	-0.1449
v(s,a)	6	7	8	9	10	
$a = a_R$	-1.3735	-2.7936	-4.3802	-6.1148	-7.9846	
$a = a_A$	-1.7936	-3.3802	-5.1148	-7.1148	-10.0705	
$v(s)$	-1.3735	-2.7936	-4.3802	-6.1148	-7.9846	

Table 3.7: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	1.9999	1.9999	1.9999	1.7774	1.3569	0.76
$a = a_A$	2.9999	2.9999	2.7774	2.3569	1.76	1.0054
$v(s)$	2.9999	2.9999	2.7774	2.3569	1.76	1.0054
v(s,a)	6	7	8	9	10	
$a = a_R$	0.0054	-0.8904	-1.9133	-3.0508	-4.2926	
$a = a_A$	0.1096	-0.9133	-2.0508	-4.0508	-7.1976	
$v(s)$	0.1096	-0.8904	-1.9133	-3.0508	-4.2926	

Next, the value of α is changed to $\alpha = 0.2$, and the values of $V(s,a)$ as shown in Table 3.8.

Table 3.8: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	0.9993	0.9993	0.9986	0.6621	0.098	-0.6241
$a = a_A$	1.9986	1.9986	1.6621	1.098	0.3759	-0.4597
$v(s)$	1.9986	1.9986	1.6621	1.098	0.3759	-0.4597
v(s,a)	6	7	8	9	10	
$a = a_R$	-1.4597	-2.3816	-3.3731	-4.4241	-5.5284	
$a = a_A$	-1.3816	-2.3731	-3.4241	-5.4241	-7.0103	
$v(s)$	-1.3816	-2.3731	-3.3731	-4.4241	-5.5284	

From these tables we see that both $v(s, a_R)$ and $v(s, a_A)$ are concave nonincreasing on s, so the deterministic stationary decision rule of optimal policy is and the optimal policy is a control limit policy. All these results testify our theorems.

Next, if we the cost function as $f() = 1$ constantly, which does not fit the theorem requiremnt, the values of $V(s,a)$ are shown in Table 3.9.

Table 3.9: Total Expected Discounted Reward $V(s,a)$

v(s,a)	s=0	1	2	3	4	5
$a = a_R$	0	0	-2.0017	-3.6105	-4.9156	-5.9919
$a = a_A$	-1.0017	-1.0017	-2.6105	-3.9156	-4.9919	-5.9005
$v(s)$	0	0	-2.0017	-3.6105	-4.9156	-5.9005
v(s,a)	6	7	8	9	10	
$a = a_R$	-6.9005	-7.6891	-8.3937	-9.0409	-9.6479	
$a = a_A$	-6.6891	-7.3937	-8.0409	-10.0409	-9.1144	
$v(s)$	-6.6891	-7.3937	-8.0409	-9.0409	-9.1144	

As seen from Table 3.9, there is no limit in the decision rule, so the optimal policy is not a control limit policy.

3.3 Non Priority Schemes

From this section we separate the new calls and handoff calls and treat them as two classes. Both classes have their own characteristics. In this section we study the performance of NPS schemes.

3.3.1 Model Formulation

To study on the NPS schemes, we make the following assumptions:

1. The arrival process of new calls and handoff calls are Poisson processes with rates λ_n and λ_h.

2. There are C channels in the system. The channel holding time for new calls and handoff calls follows negative exponential distribution with rate μ_n and μ_h.

3. Accepting an incoming new(handoff) call would contribute $R_n(R_h)$ units of reward to the system. If the number of new calls and handoff calls are s_n and s_h, the system incurs a holding cost rate $f(s_n, s_h)$ per unit time.

For our admission problem, since there are two classes of calls, a policy is called a control limit policy if for a given s_n (the number of new calls), there exists a constant or threshold $N \geq 0$ for accepting the handoff calls, that is

$$d(s_n, s_h, A_h) = \begin{cases} Admit, & s_h \leq N, \\ Reject, & s_h > N. \end{cases}$$

Similar definition could be used on $d(s_n, s_h, A_n)$ when s_h is given. It can be seen that a control limit policy is a stationary deterministic policy.

As we know, each SMDP model can be uniquely identified by the following five components: the decision epochs, the state space, the action space, the reward function, and the transition probabilities. Suppose there are K reserved channels for handoff calls, the SMDP models for RCS and NPS schemes could be defined as follows:

1. Let the state variable consists of the number of new calls and handoff calls in the system, the status of calls leaving, arriving to the system. For the NPS scheme, $S = \{0, 1, \ldots, C\} \times \{0, 1, \ldots, C\} \times \{D, A_n, A_h\}$, $s_n + s_h \leq C$. Where in the third set D means a departure from the system, while A_n means an arrival of a new call and A_h is an arrival of a handoff call. In general a state could be written as $\langle s_n, s_h, b\rangle$, where s_n and s_h are the numbers of new calls and handoff calls, and b stands for the last call event, $b \in \{D, A_n, A_h\}$.

2. For states $\langle s_n, s_h, D\rangle$, set a_C as the action to continue, thus $A_{\langle s_n, s_h, D\rangle} = \{a_C\}$. In states $\langle s_n, s_h, A_n\rangle$ and $\langle s_n, s_h, A_h\rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle s_n, s_h, A_n\rangle} = A_{\langle s_n, s_h, A_h\rangle} = \{a_R, a_A\}$.

3. For both schemes, the decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision

epochs starting from state s be

$$F(t|s,a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle s_n, s_h, b\rangle$ and action a, let $\beta_0 = \lambda_n + \lambda_h + s_n\mu_n + s_h\mu_h$, so $\beta(s,a)$ can be written as

$$\beta(\langle s_n, s_h, b\rangle, a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + \mu_n, & a = a_A, b = A_n, \\ \beta_0 + \mu_h, & a = a_A, b = A_h. \end{cases}$$

4. Let $q(j|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle s_n, s_h, b\rangle$, $b = D, a = a_C$ and $b = \{A_n, A_h\}, a = a_R$, let $\bar{s_n} = \max(s_n - 1, 0)$, $\bar{s_h} = \max(s_h - 1, 0)$, the state transition probabilities are

$$q(j|s,a) = \begin{cases} \lambda_n/\beta_0, & j = \langle s_n, s_h, A_n\rangle, \\ \lambda_h/\beta_0, & j = \langle s_n, s_h, A_h\rangle, \\ s_n\mu_n/\beta_0, & j = \langle \bar{s_n}, s_h, D\rangle, \\ s_h\mu_h/\beta_0, & j = \langle s_n, \bar{s_h}, D\rangle. \end{cases}$$

And, for states $s = \langle s_n, s_h, b\rangle$, $b = \{A_n, A_h\}$ and $a = a_A$, since admitting an incoming call migrates the system state immediately, there is

$$q(j|s,a_A) = \begin{cases} q(j|\langle s_n+1, s_h, D\rangle, a_R), & b = A_n, \\ q(j|\langle s_n, s_h+1, D\rangle, a_R), & b = A_h. \end{cases}$$

5. Because the system state does not change between decision epochs, the expected

discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a) E_s^a \left\{ \int_0^t e^{-\alpha\tau} d\tau \right\} \\ &= k(s,a) + c(s,a) E_s^a \left\{ [1 - e^{-\alpha t}]/\alpha \right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha + \beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \text{any } b, \\ R_n, & a = a_A, b = A_n, \\ R_h, & a = a_A, b = A_h, \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(s_n, s_h), & a = a_C, a = a_R, \text{any } b, \\ -f(s_n + 1, s_h), & a = a_A, b = A_n, \\ -f(s_n, s_h + 1), & a = a_A, b = A_h. \end{cases}$$

Based on our assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so our problem can be reduced to as finding a deterministic decision rule d. For each deterministic decision rule d, let $q_d(j|s) = q(j|s, d(s))$, $r_d(s) = r(s, d(s))$ and $\beta_d(s) = \beta(s, d(s))$, we have

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{j \in S} q_d(j|s) v_\alpha^{d^\infty}(j). \tag{3.6}$$

From equation (3.6), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1 - q(s|s,a)]\beta(s,a) \leq c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = \lambda_n + \lambda_h + C * \max(\mu_n, \mu_h)$, we have

$$\tilde{q}(j|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & j = s, \\ \frac{q(j|s,a)\beta(s,a)}{c}, & j \neq s. \end{cases}$$

So, for states $s = \langle s_n, s_h, D\rangle$ and $a = a_C$, we have

$$\tilde{q}(j|s,a_C) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h, A_n\rangle, \\ \lambda_h/c, & j = \langle s_n, s_h, A_h\rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D\rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D\rangle \\ (c-\beta_0)/c, & j = s. \end{cases}$$

And, for states $s = \langle s_n, s_h, A_n\rangle$ and $a = a_R$, we get

$$\tilde{q}(j|s,a_R) = \begin{cases} (c+\lambda_n-\beta_0)/c, & j = \langle s_n, s_h, A_n\rangle, \\ \lambda_h/c, & j = \langle s_n, s_h, A_h\rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D\rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D\rangle. \end{cases}$$

Also, for states $s = \langle s_n, s_h, A_h \rangle$ and $a = a_R$,

$$\tilde{q}(j|s, a_R) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h, A_n \rangle, \\ (c + \lambda_h - \beta_0)/c, & j = \langle s_n, s_h, A_h \rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D \rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D \rangle. \end{cases}$$

Similarly, for states $s = \langle s_n, s_h, A_n \rangle$ and $a = a_A$,

$$\tilde{q}(j|s, a_A) = \begin{cases} \lambda_n/c, & j = \langle s_n + 1, s_h, 1 \rangle, \\ \lambda_h/c, & j = \langle s_n + 1, s_h, 2 \rangle, \\ (s_n + 1)\mu_n/c, & j = \langle s_n, s_h, D \rangle, \\ s_h\mu_h/c, & j = \langle s_n + 1, \bar{s_h}, D \rangle, \\ (c - \beta_0 - \mu_n)/c, & j = s. \end{cases}$$

Finally, for states $s = \langle s_n, s_h, A_h \rangle$ and $a = a_A$,

$$\tilde{q}(j|s, a_A) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h + 1, 1 \rangle, \\ \lambda_h/c, & j = \langle s_n, s_h + 1, 2 \rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h + 1, D \rangle, \\ (s_h + 1)\mu_h/c, & j = \langle s_n, s_h, D \rangle, \\ (c - \beta_0 - \mu_h)/c, & j = s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s, a) \equiv r(s, a)\frac{\alpha + \beta(s, a)}{\alpha + c}.$$

For each d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{3.7}$$

From equation (3.6) and (3.7), the optimality equation of $v(s)$ for maximum $v_\alpha^\pi(s)$ would have the form of

$$v(s) = \max_{a \in A_s} \left\{ \tilde{r}(s,a) + \lambda \sum_{j \in S} \tilde{q}(j|s,a) v(j) \right\}, \tag{3.8}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (3.8), we have for states with $b = D$,

$$\begin{aligned} v(\langle s_n, s_h, D\rangle) \quad = \quad & \frac{1}{\alpha + c}[s_n\mu_n v(\langle \bar{s_n}, s_h, D\rangle) + \lambda_n v(\langle s_n, s_h, A_n\rangle) + \lambda_h v(\langle s_n, s_h, A_h\rangle) \\ & + s_h\mu_h v(\langle s_n, \bar{s_h}, D\rangle) + (c - \beta_0) v(\langle s_n, s_h, D\rangle) - f(s_n, s_h)]. \end{aligned} \tag{3.9}$$

After rearranging the terms, equation (3.9) can be rewritten as

$$\begin{aligned} v(\langle s_n, s_h, D\rangle) \quad = \quad & \frac{1}{\alpha + \beta_0}[-f(s_n, s_h) + \lambda_n v(\langle s_n, s_h, A_n\rangle) + \lambda_h v(\langle s_n, s_h, A_h\rangle) \\ & + s_n\mu_n v(\langle \bar{s_n}, s_h, D\rangle) + s_h\mu_h v(\langle s_n, \bar{s_h}, D\rangle)]. \end{aligned}$$

3.3.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_n$,

$$v(\langle s_n, s_h, A_n\rangle) = \max\left[v(\langle s_n, s_h, D\rangle), R_n + v(\langle s_n + 1, s_h, D\rangle)\right].$$

Similarly, for states $b = A_h$,

$$v(\langle s_n, s_h, A_h\rangle) = \max\left[v(\langle s_n, s_h, D\rangle), R_h + v(\langle s_n, s_h + 1, D\rangle)\right].$$

Theorem 3.4 *If $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on both s_n and s_h, the optimal policy for both RCS and NPS schemes is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_h(\langle s_n, s_h, D\rangle) = v(\langle s_n, s_h + 1, D\rangle) - v(\langle s_n, s_h, D\rangle)$, so $\triangle v_h(\langle s_n, s_h, D\rangle)$ is ≤ 0 and is nonincreasing, for states with fixed s_n, we have the decision rule

$$d(s_n, s_h, A_h) = \begin{cases} a_A, & \triangle v_h(\langle s_n, s_h, D\rangle) > -R_h, \\ a_R, & \triangle v_h(\langle s_n, s_h, D\rangle) \leq -R_h. \end{cases}$$

So, if $d(s_n, s_h, A_h) = a_R$, we have $d(s_n, s_h+1, A_h) = a_R$, and so on. Similarly for states with fixed s_h, if $d(s_n, s_h, A_n) = a_R$, we have $d(s_n + 1, s_h, A_n) = a_R$. Consequently the optimal policy for both RCS and NPS schemes is a control limit policy (or threshold policy).

Theorem 3.5 *Suppose that the cost function $f(i, j)$ is convex and nondecreasing on both i and j, and*

$$f(i, j) + f(i + 1, j + 1) \geq f(i + 1, j) + f(i, j + 1),$$

then (1): $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on both s_n and s_h, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (3.9) to obtain $v^1(\langle s_n, s_h, D\rangle) = \frac{-f(s_n,s_h)}{\alpha+c}$.

2. Set n=1. We have $v^n(\langle s_n, s_h, D\rangle)$ which is concave nonincreasing. From Lemma A.1 and A.3, $v^n(\langle s_n, s_h, A_n\rangle)$ and $v^n(\langle s_n, s_h, A_h\rangle)$ are also concave nonincreasing

on both s_n and s_h. And from Lemma A.2,

$$v^n(\langle s_n, s_h, b\rangle) + v^n(\langle s_n+1, s_h+1, b\rangle) \leq v^n(\langle s_n+1, s_h, b\rangle) + v^n(\langle s_n, s_h+1, b\rangle),$$

for $b = A_n, A_h$.

3. Substituting these v^n back into equation (3.9), we get

$$\begin{aligned} v^{n+1}(\langle s_n, s_h, D\rangle) &= \frac{1}{\alpha+c}[-f(s_n, s_h) + \lambda_n v^n(\langle s_n, s_h, A1\rangle) + \lambda_h v^n(\langle s_n, s_h, A2\rangle) \\ &\quad + s_n\mu_n v^n(\langle \bar{s_n}, s_h, D\rangle) + s_h\mu_h v^n(\langle s_n, \bar{s_h}, D\rangle) \\ &\quad +(c-\beta)v^n(\langle s_n, s_h, D\rangle)], \end{aligned}$$

so we have

$$\begin{aligned} & v^{n+1}(\langle s_n+1, s_h, D\rangle) - v^{n+1}(\langle s_n, s_h, D\rangle) \\ &= \frac{1}{\alpha+c}[(f(s_n, s_h) - f(s_n+1, s_h)) \\ &+\lambda_n(v^n(\langle s_n+1, s_h, A_n\rangle) - v^n(\langle s_n, s_h, A_n\rangle)) \\ &+\lambda_h(v^n(\langle s_n, s_h+1, A_h\rangle) - v^n(\langle s_n, s_h, A_h\rangle)) \\ &+s_n\mu_n(v^n(\langle s_n, s_h, D\rangle) - v^n(\langle \bar{s_n}, s_h, D\rangle)) \\ &+s_h\mu_h(v^n(\langle s_n+1, \bar{s_h}, D\rangle) - v^n(\langle s_n, \bar{s_h}, D\rangle)) \\ &+c_1(v^n(\langle s_n+1, s_h, D\rangle) - v^n(\langle s_n, s_h, D\rangle))], \end{aligned}$$

where $c_1 = c - \beta_0 - \mu_n$. Since $v^n(\langle s_n, s_h, b\rangle)$, $b = D, A_n, A_h$ are concave nonincreasing on s_n, the combination of these functions $v^{n+1}(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on s_n. Similarly taking the same process for $v^{n+1}(\langle s_n, s_h+1, D\rangle) - v^{n+1}(\langle s_n, s_h, D\rangle)$, then $v^{n+1}(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on s_h. Also, using equation (3.9),

$$v^{n+1}(\langle s_n, s_h, D\rangle) + v^{n+1}(\langle s_n+1, s_h+1, D\rangle)$$
$$-v^{n+1}(\langle s_n+1, s_h, D\rangle) - v^{n+1}(\langle s_n, s_h+1, D\rangle),$$

could be divided into several different parts of $v^n(\langle s_n, s_h, b\rangle) + v^n(\langle s_n+1, s_h+1, b\rangle) - v^n(\langle s_n+1, s_h, b\rangle) - v^n(\langle s_n, s_h+1, b\rangle), b = D, A_n, A_h$, which are all ≤ 0. So

$$v^{n+1}(\langle s_n, s_h, D\rangle) + v^{n+1}(\langle s_n+1, s_h+1, D\rangle)$$
$$\leq v^{n+1}(\langle s_n+1, s_h, D\rangle) + v^{n+1}(\langle s_n, s_h+1, D\rangle).$$

4. Set n=n+1, go back to step 2.

5. As the iteration continues, $v^n(\langle s_n, s_h, D\rangle)$ is concave nonincreasing, the solution of $v(s)$ is unique, so the value iteration $v^n(s)$ converges to $v(s)$, so $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing.

6. From Theorem 3.4, the optimal policy is a control limit policy. The verification is over.

3.3.3 Simulation Results

Let the cost function be $f(s_n, s_h) = s_n + s_h$, which fits for the theorem requirement. We set the parameters for simulation as in Table 3.15 to study the performance of the NPS schemes.

As shown in Table 3.15, the new calls and handoff calls have different arrival rates, service rates and rewards, but they have the same holding cost rate, which fits for the property of cost functions in Theorem 3.5. For both types of calls, the arrival

Table 3.10: Parameter Values

Parameters	λ_n	λ_h	μ_n	μ_h	R_n	R_h	C	α
Values	4	2	6	4	0.164	0.3	10	0.1

rates are smaller than the service rates, which represent a light traffic load. With the parameters setting in Table 3.10, using the Value Iteration Method to solve the equations, we get the actions for each state as shown in Table 3.11. Because of the paper size, for values of $v(s)$, only those states $b = D$ are listed.

Table 3.11: Actions for call arrivals

b=A_n	$s_h \rightarrow$										
	1	1	1	1	1	1	1	1	0	0	0
	1	1	1	1	1	1	1	0	0	0	-1
	1	1	1	1	1	1	0	0	0	-1	-1
	1	1	1	1	1	0	0	0	-1	-1	-1
$s_n \downarrow$	1	1	1	1	0	0	0	-1	-1	-1	-1
	1	1	1	0	0	0	-1	-1	-1	-1	-1
	1	1	0	0	0	-1	-1	-1	-1	-1	-1
	1	0	0	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
b=A_h	$s_h \rightarrow$										
	1	1	1	1	1	1	1	1	1	1	0
	1	1	1	1	1	1	1	1	1	0	-1
	1	1	1	1	1	1	1	1	0	-1	-1
	1	1	1	1	1	1	1	0	-1	-1	-1
	1	1	1	1	1	1	0	-1	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	0	-1	-1	-1	-1	-1
	1	1	1	1	0	-1	-1	-1	-1	-1	-1
	1	1	1	0	-1	-1	-1	-1	-1	-1	-1
	1	1	0	-1	-1	-1	-1	-1	-1	-1	-1
	1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

It is seen that in Table 3.11 '1' means to accept the call and '0' means to reject the call on such state, and '-1' means that this state does not exist in the state space.

Similarly, we get the values of $v(\langle s_n, s_h, D\rangle)$ for each state as shown in Table 3.12.

Table 3.12: Total discounted expected reward $v(s)$

b=D	s_h=0	1	2	3	4	5
	1.1245	0.8806	0.6367	0.3928	0.1489	-0.095
	0.9606	0.7167	0.4728	0.2289	-0.015	-0.2589
	0.7966	0.5527	0.3088	0.0649	-0.179	-0.4229
	0.6327	0.3888	0.1449	-0.099	-0.3429	-0.5868
	0.4688	0.2249	-0.019	-0.2629	-0.5069	-0.7509
$s_n \downarrow$	0.3048	0.0609	-0.183	-0.4269	-0.6709	-0.9171
	0.1409	-0.103	-0.3469	-0.5909	-0.837	0
	-0.023	-0.267	-0.5109	-0.7569	0	0
	-0.187	-0.431	-0.6769	0	0	0
	-0.351	-0.5968	0	0	0	0
	-0.5168	0	0	0	0	0
	6	7	8	9	10	
	-0.3389	-0.5828	-0.8267	-1.0708	-1.3175	
	-0.5029	-0.7468	-0.9908	-1.2374	0	
	-0.6668	-0.9108	-1.1573	0	0	
	-0.8309	-1.0772	0	0	0	
	-0.9971	0	0	0	0	
$s_n \downarrow$	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	

And, we can see that the values of $\triangle v(s)$ in Table 3.12 are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion.

If we change λ_n=24, λ_h=12 and keep all the parameters the same as in Table 3.10, this would make the system in a heavier traffic load, the actions and values of $v(\langle s_n, s_h, D\rangle)$ are shown in Table 3.13 and Table 3.14. As seen from Table 3.13,

Table 3.13: Actions for call arrivals

b=A_n	$s_h \rightarrow$									
	1	1	1	0	0	0	0	0	0	0	0
	1	1	1	0	0	0	0	0	0	0	-1
	1	1	0	0	0	0	0	0	0	-1	-1
	1	0	0	0	0	0	0	0	-1	-1	-1
$s_n \downarrow$	0	0	0	0	0	0	0	-1	-1	-1	-1
	0	0	0	0	0	0	-1	-1	-1	-1	-1
	0	0	0	0	0	-1	-1	-1	-1	-1	-1
	0	0	0	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
b=A_h	$s_h \rightarrow$										
	1	1	1	1	1	1	1	1	1	1	0
	1	1	1	1	1	1	1	1	1	0	-1
	1	1	1	1	1	1	1	1	0	-1	-1
	1	1	1	1	1	1	1	0	-1	-1	-1
	1	1	1	1	1	1	0	-1	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	0	-1	-1	-1	-1	-1
	1	1	1	1	0	-1	-1	-1	-1	-1	-1
	1	1	1	0	-1	-1	-1	-1	-1	-1	-1
	1	1	0	-1	-1	-1	-1	-1	-1	-1	-1
	1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

compared to Table 3.11, the limit to accept the new call is smaller, this is because the system is in a heavier traffic load.

Table 3.14: Total discounted expected reward $v(s)$

b=D	s_h=0	1	2	3	4	5
	6.7284	6.4844	6.2404	5.9964	5.7523	5.5082
	6.5644	6.3204	6.0764	5.8324	5.5883	5.344
	6.4005	6.1565	5.9124	5.6683	5.4241	5.1796
	6.2365	5.9925	5.7484	5.5042	5.2598	5.0144
	6.0725	5.8284	5.5843	5.3399	5.0947	4.8467
$s_n \downarrow$	5.9085	5.6644	5.42	5.175	4.9273	4.6696
	5.7444	5.5001	5.2552	5.0078	4.7506	0
	5.5802	5.3354	5.0882	4.8315	0	0
	5.4156	5.1686	4.9124	0	0	0
	5.249	4.9931	0	0	0	0
	5.0739	0	0	0	0	0
	6	7	8	9	10	
	5.2639	5.0191	4.7732	4.5237	4.263	
	5.0994	4.8536	4.6046	4.3446	0	
	4.934	4.6854	4.426	0	0	
	4.7661	4.5073	0	0	0	
	4.5885	0	0	0	0	
$s_n \downarrow$	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	

We can see that the values of $\triangle v(s)$ in Table 3.14 are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion. Compared to Table 3.12, the values of $v(\langle s_n, s_h, D\rangle)$ is larger, this is because the system is in a heavier traffic load.

3.4 Reserved Channel Schemes

Since it is always more desirable not to drop a handoff call than a new call, in this section we separate the new calls and handoff calls. There are some channel allocation schemes for accepting new calls and handoff calls. In the RCS scheme, some channels are reserved for the handoff calls. A handoff call can only be blocked if all the channels are busy, while a new call can be blocked even when there are several free channels. In a word, the handoff calls have some priority over the new calls.

3.4.1 Model Formulation

To study on the RCS schemes, we make the following assumptions:

1. The arrival process of new calls and handoff calls are Poisson processes with rates λ_n and λ_h.

2. There are C channels in the system. The channel holding time for new calls and handoff calls follows negative exponential distribution with rate μ_n and μ_h.

3. Accepting an incoming new(handoff) call would contribute $R_n(R_h)$ units of reward to the system. If the number of new calls and handoff calls are s_n and s_h, the system incurs a holding cost rate $f(s_n, s_h)$ per unit time.

4. There are K channels reserved for the handoff calls.

For our admission problem, since there are two classes of calls, a policy is called a control limit policy if for a given s_n (the number of new calls), there exists a constant or threshold $N \geq 0$ for accepting the handoff calls, that is

$$d(s_n, s_h, A_h) = \begin{cases} Admit, & s_h \leq N, \\ Reject, & s_h > N. \end{cases}$$

Similar definition could be used on $d(s_n, s_h, A_n)$ when s_h is given. It can be seen that a control limit policy is a stationary deterministic policy.

As we know, each SMDP model can be uniquely identified by the following five components: the decision epochs, the state space, the action space, the reward function, and the transition probabilities. Suppose there are K reserved channels for handoff calls, the SMDP models for RCS schemes could be defined as follows:

1. Let the state variable consists of the number of new calls and handoff calls in the system, the status of calls leaving, arriving to the system. For the RCS scheme, the state space is $S = \{0, 1, \ldots, C-K\} \times \{0, 1, \ldots, C\} \times \{D, A_n, A_h\}$, $s_n \le \max(C - K - s_h, 0)$. Where in the third set D means a departure from the system, while A_n means an arrival of a new call and A_h is an arrival of a handoff call. In general a state could be written as $\langle s_n, s_h, b\rangle$, where s_n and s_h are the numbers of new calls and handoff calls, and b stands for the last call event, $b \in \{D, A_n, A_h\}$.

2. In both schemes, for states $\langle s_n, s_h, D\rangle$, set a_C as the action to continue, thus $A_{\langle s_n, s_h, D\rangle} = \{a_C\}$. In states $\langle s_n, s_h, A_n\rangle$ and $\langle s_n, s_h, A_h\rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle s_n, s_h, A_n\rangle} = A_{\langle s_n, s_h, A_h\rangle} = \{a_R, a_A\}$.

3. For both schemes, the decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s, a) = 1 - e^{-\beta(s,a)t}, t \ge 0.$$

For each state $s = \langle s_n, s_h, b\rangle$ and action a, let $\beta_0 = \lambda_n + \lambda_h + s_n\mu_n + s_h\mu_h$, so

$\beta(s,a)$ can be written as

$$\beta(\langle s_n, s_h, b\rangle, a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + \mu_n, & a = a_A, b = A_n, \\ \beta_0 + \mu_h, & a = a_A, b = A_h. \end{cases}$$

4. For both schemes, let $q(j|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle s_n, s_h, b\rangle$, $b = D, a = a_C$ and $b = \{A_n, A_h\}, a = a_R$, let $\bar{s_n} = \max(s_n - 1, 0)$, $\bar{s_h} = \max(s_h - 1, 0)$, the state transition probabilities are

$$q(j|s,a) = \begin{cases} \lambda_n/\beta_0, & j = \langle s_n, s_h, A_n\rangle, \\ \lambda_h/\beta_0, & j = \langle s_n, s_h, A_h\rangle, \\ s_n\mu_n/\beta_0, & j = \langle \bar{s_n}, s_h, D\rangle, \\ s_h\mu_h/\beta_0, & j = \langle s_n, \bar{s_h}, D\rangle. \end{cases}$$

And, for states $s = \langle s_n, s_h, b\rangle$, $b = \{A_n, A_h\}$ and $a = a_A$, since admitting an incoming call migrates the system state immediately, there is

$$q(j|s,a_A) = \begin{cases} q(j|\langle s_n+1, s_h, D\rangle, a_R), & b = A_n, \\ q(j|\langle s_n, s_h+1, D\rangle, a_R), & b = A_h. \end{cases}$$

5. For both schemes, because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a)E_s^a\left\{\int_0^t e^{-\alpha\tau}d\tau\right\} \\ &= k(s,a) + c(s,a)E_s^a\left\{[1-e^{-\alpha t}]/\alpha\right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha+\beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \text{any } b, \\ R_n, & a = a_A, b = A_n, \\ R_h, & a = a_A, b = A_h, \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(s_n, s_h), & a = a_C, a = a_R, \text{any } b, \\ -f(s_n+1, s_h), & a = a_A, b = A_n, \\ -f(s_n, s_h+1), & a = a_A, b = A_h. \end{cases}$$

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so our problem can be reduced to as finding a deterministic decision rule d. For each deterministic decision rule d, let $q_d(j|s) = q(j|s, d(s))$, $r_d(s) = r(s, d(s))$ and $\beta_d(s) = \beta(s, d(s))$, we have

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{j \in S} q_d(j|s) v_\alpha^{d^\infty}(j). \tag{3.10}$$

From equation (3.10), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1 - q(s|s,a)]\beta(s,a) \le c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = \lambda_n + \lambda_h + C * \max(\mu_n, \mu_h)$, we have

$$\tilde{q}(j|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & j = s, \\ \frac{q(j|s,a)\beta(s,a)}{c}, & j \neq s. \end{cases}$$

So, for states $s = \langle s_n, s_h, D \rangle$ and $a = a_C$, we have

$$\tilde{q}(j|s,a_C) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h, A_n \rangle, \\ \lambda_h/c, & j = \langle s_n, s_h, A_h \rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D \rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D \rangle \\ (c-\beta_0)/c, & j = s. \end{cases}$$

And, for states $s = \langle s_n, s_h, A_n \rangle$ and $a = a_R$, we get

$$\tilde{q}(j|s,a_R) = \begin{cases} (c+\lambda_n-\beta_0)/c, & j = \langle s_n, s_h, A_n \rangle, \\ \lambda_h/c, & j = \langle s_n, s_h, A_h \rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D \rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D \rangle. \end{cases}$$

Also, for states $s = \langle s_n, s_h, A_h \rangle$ and $a = a_R$,

$$\tilde{q}(j|s,a_R) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h, A_n \rangle, \\ (c+\lambda_h-\beta_0)/c, & j = \langle s_n, s_h, A_h \rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h, D \rangle, \\ s_h\mu_h/c, & j = \langle s_n, \bar{s_h}, D \rangle. \end{cases}$$

Similarly, for states $s = \langle s_n, s_h, A_n\rangle$ and $a = a_A$,

$$\tilde{q}(j|s,a_A) = \begin{cases} \lambda_n/c, & j = \langle s_n+1, s_h, 1\rangle, \\ \lambda_h/c, & j = \langle s_n+1, s_h, 2\rangle, \\ (s_n+1)\mu_n/c, & j = \langle s_n, s_h, D\rangle, \\ s_h\mu_h/c, & j = \langle s_n+1, \bar{s_h}, D\rangle, \\ (c-\beta_0-\mu_n)/c, & j = s. \end{cases}$$

Finally, for states $s = \langle s_n, s_h, A_h\rangle$ and $a = a_A$,

$$\tilde{q}(j|s,a_A) = \begin{cases} \lambda_n/c, & j = \langle s_n, s_h+1, 1\rangle, \\ \lambda_h/c, & j = \langle s_n, s_h+1, 2\rangle, \\ s_n\mu_n/c, & j = \langle \bar{s_n}, s_h+1, D\rangle, \\ (s_h+1)\mu_h/c, & j = \langle s_n, s_h, D\rangle, \\ (c-\beta_0-\mu_h)/c, & j = s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha+\beta(s,a)}{\alpha+c}.$$

For each d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{3.11}$$

From equation (3.10) and (3.11), the optimality equation of $v(s)$ for maximum $v_\alpha^\pi(s)$ would have the form of

$$v(s) = \max_{a\in A_s}\left\{\tilde{r}(s,a) + \lambda\sum_{j\in S}\tilde{q}(j|s,a)v(j)\right\}, \tag{3.12}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (3.12), we have for states with $b = D$,

$$\begin{aligned} v(\langle s_n, s_h, D\rangle) &= \frac{1}{\alpha + c}[s_n\mu_n v(\langle \bar{s_n}, s_h, D\rangle) + \lambda_n v(\langle s_n, s_h, A_n\rangle) + \lambda_h v(\langle s_n, s_h, A_h\rangle) \\ &\quad + s_h\mu_h v(\langle s_n, \bar{s_h}, D\rangle) + (c - \beta_0)v(\langle s_n, s_h, D\rangle) - f(s_n, s_h)]. \end{aligned} \tag{3.13}$$

3.4.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_n$,

$$v(\langle s_n, s_h, A_n\rangle) = \max\left[v(\langle s_n, s_h, D\rangle), R_n + v(\langle s_n + 1, s_h, D\rangle)\right].$$

Similarly, for states $b = A_h$,

$$v(\langle s_n, s_h, A_h\rangle) = \max\left[v(\langle s_n, s_h, D\rangle), R_h + v(\langle s_n, s_h + 1, D\rangle)\right].$$

Theorem 3.6 *If $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on both s_n and s_h, the optimal policy for both RCS and NPS schemes is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_h(\langle s_n, s_h, D\rangle) = v(\langle s_n, s_h + 1, D\rangle) - v(\langle s_n, s_h, D\rangle)$, so $\triangle v_h(\langle s_n, s_h, D\rangle)$ is ≤ 0 and is nonincreasing, for states with fixed s_n, we have the decision rule

$$d(s_n, s_h, A_h) = \begin{cases} a_A, & \triangle v_h(\langle s_n, s_h, D\rangle) > -R_h, \\ a_R, & \triangle v_h(\langle s_n, s_h, D\rangle) \leq -R_h. \end{cases}$$

So, if $d(s_n, s_h, A_h) = a_R$, we have $d(s_n, s_h+1, A_h) = a_R$, and so on. Similarly for states with fixed s_h, if $d(s_n, s_h, A_n) = a_R$, we have $d(s_n+1, s_h, A_n) = a_R$. Consequently the optimal policy for both RCS and NPS schemes is a control limit policy (or threshold policy).

Theorem 3.7 *Suppose that the cost function $f(i,j)$ is convex and nondecreasing on both i and j, and*

$$f(i,j) + f(i+1,j+1) \geq f(i+1,j) + f(i,j+1),$$

then (1): $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on both s_n and s_h, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (3.13) to obtain $v^1(\langle s_n, s_h, D\rangle) = \frac{-f(s_n,s_h)}{\alpha+c}$.

2. Set n=1. We have $v^n(\langle s_n, s_h, D\rangle)$ which is concave nonincreasing. From Lemma A.1 and A.3, $v^n(\langle s_n, s_h, A_n\rangle)$ and $v^n(\langle s_n, s_h, A_h\rangle)$ are also concave nonincreasing on both s_n and s_h. And from Lemma A.2,

   $$v^n(\langle s_n, s_h, b\rangle) + v^n(\langle s_n+1, s_h+1, b\rangle) \leq v^n(\langle s_n+1, s_h, b\rangle) + v^n(\langle s_n, s_h+1, b\rangle),$$

   for $b = A_n, A_h$.

3. Substituting these v^n back into equation (3.13), we get

   $$\begin{aligned} v^{n+1}(\langle s_n, s_h, D\rangle) \;&=\; \frac{1}{\alpha+c}[-f(s_n,s_h) + \lambda_n v^n(\langle s_n, s_h, A1\rangle) + \lambda_h v^n(\langle s_n, s_h, A2\rangle) \\ &\quad + s_n\mu_n v^n(\langle \bar{s_n}, s_h, D\rangle) + s_h\mu_h v^n(\langle s_n, \bar{s_h}, D\rangle) \\ &\quad + (c-\beta)v^n(\langle s_n, s_h, D\rangle)], \end{aligned}$$

so we have

$$
\begin{aligned}
&v^{n+1}(\langle s_n+1, s_h, D\rangle) - v^{n+1}(\langle s_n, s_h, D\rangle) \\
&= \frac{1}{\alpha+c}[(f(s_n, s_h) - f(s_n+1, s_h)) \\
&+\lambda_n(v^n(\langle s_n+1, s_h, A_n\rangle) - v^n(\langle s_n, s_h, A_n\rangle)) \\
&+\lambda_h(v^n(\langle s_n, s_h+1, A_h\rangle) - v^n(\langle s_n, s_h, A_h\rangle)) \\
&+s_n\mu_n(v^n(\langle s_n, s_h, D\rangle) - v^n(\langle \bar{s_n}, s_h, D\rangle)) \\
&+s_h\mu_h(v^n(\langle s_n+1, \bar{s_h}, D\rangle) - v^n(\langle s_n, \bar{s_h}, D\rangle)) \\
&+c_1(v^n(\langle s_n+1, s_h, D\rangle) - v^n(\langle s_n, s_h, D\rangle))],
\end{aligned}
$$

where $c_1 = c - \beta_0 - \mu_n$. Since $v^n(\langle s_n, s_h, b\rangle)$, $b = D, A_n, A_h$ are concave nonincreasing on s_n, the combination of these functions $v^{n+1}(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on s_n. Similarly taking the same process for $v^{n+1}(\langle s_n, s_h+1, D\rangle) - v^{n+1}(\langle s_n, s_h, D\rangle)$, then $v^{n+1}(\langle s_n, s_h, D\rangle)$ is concave nonincreasing on s_h. Also, using equation (3.13),

$$
\begin{aligned}
&v^{n+1}(\langle s_n, s_h, D\rangle) + v^{n+1}(\langle s_n+1, s_h+1, D\rangle) \\
&-v^{n+1}(\langle s_n+1, s_h, D\rangle) - v^{n+1}(\langle s_n, s_h+1, D\rangle),
\end{aligned}
$$

could be divided into several different parts of $v^n(\langle s_n, s_h, b\rangle) + v^n(\langle s_n+1, s_h+1, b\rangle) - v^n(\langle s_n+1, s_h, b\rangle) - v^n(\langle s_n, s_h+1, b\rangle), b = D, A_n, A_h$, which are all ≤ 0. So

$$
\begin{aligned}
&v^{n+1}(\langle s_n, s_h, D\rangle) + v^{n+1}(\langle s_n+1, s_h+1, D\rangle) \\
&\leq v^{n+1}(\langle s_n+1, s_h, D\rangle) + v^{n+1}(\langle s_n, s_h+1, D\rangle).
\end{aligned}
$$

4. Set n=n+1, go back to step 2.

5. As the iteration continues, $v^n(\langle s_n, s_h, D\rangle)$ is concave nonincreasing, the solution of $v(s)$ is unique, so the value iteration $v^n(s)$ converges to $v(s)$, so $v(\langle s_n, s_h, D\rangle)$ is concave nonincreasing.

6. From Theorem 3.6, the optimal policy is a control limit policy. The verification is over.

3.4.3 Simulation Results

Let the reserved channels for handoff calls be $K = 2$ and cost function be $f(s_n, s_h) = s_n + s_h$, with x, y be some constants. We set the parameters for simulation as in Table 3.15 (same as Table 3.10) to study the performance of the RCS.

Table 3.15: Simulation Parameter Values

Parameters	λ_n	λ_h	μ_n	μ_h	R_n	R_h	C	α
Values	4	2	6	4	0.164	0.3	10	0.1

As shown in Table 3.15, the new calls and handoff calls have different arrival rates, service rates and rewards, but they have the same holding cost rate, which fits for the property of cost functions in Theorem 3.7. For both types of calls, the arrival rates are smaller than the service rates, which represent a light traffic load. With the parameters setting in Table 3.15, using the Value Iteration Method to solve the equations, we get the actions for each state as shown in Table 3.16. Because of the paper size, for values of $v(s)$, only those states $b = D$ are listed.

Table 3.16: Actions for call arrivals

b=A_n	$s_h \rightarrow$									
	1	1	1	1	1	1	1	1	0	0	0
	1	1	1	1	1	1	1	0	0	0	-1
	1	1	1	1	1	1	0	0	0	-1	-1
	1	1	1	1	1	0	0	0	-1	-1	-1
$s_n \downarrow$	1	1	1	1	0	0	0	-1	-1	-1	-1
	1	1	1	0	0	0	-1	-1	-1	-1	-1
	1	1	0	0	0	-1	-1	-1	-1	-1	-1
	0	0	0	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
b=A_2	$s_h \rightarrow$										
	1	1	1	1	1	1	1	1	1	1	0
	1	1	1	1	1	1	1	1	1	0	-1
	1	1	1	1	1	1	1	1	0	-1	-1
	1	1	1	1	1	1	1	0	-1	-1	-1
	1	1	1	1	1	1	0	-1	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	0	-1	-1	-1	-1	-1
	1	1	1	1	0	-1	-1	-1	-1	-1	-1
	1	1	1	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

It is seen that in Table 3.16, '1' means to accept the call and '0' means to reject the call on such state, and '-1' means that this state does not exist in the state space.

Table 3.17: Total discounted expected reward $v(s)$

b=D	s_h=0	1	2	3	4	5
$s_n \downarrow$	1.1245	0.8806	0.6367	0.3928	0.1489	-0.095
	0.9606	0.7167	0.4728	0.2289	-0.015	-0.2589
	0.7966	0.5527	0.3088	0.0649	-0.179	-0.4229
	0.6327	0.3888	0.1449	-0.099	-0.3429	-0.5868
	0.4688	0.2249	-0.019	-0.2629	-0.5069	-0.7509
	0.3048	0.0609	-0.183	-0.4269	-0.6709	-0.9171
	0.1409	-0.103	-0.3469	-0.5909	-0.837	0
	-0.023	-0.267	-0.5109	-0.7569	0	0
	-0.1893	-0.4332	-0.6772	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	6	7	8	9	10	
$s_n \downarrow$	-0.3389	-0.5828	-0.8267	-1.0708	-1.3175	
	-0.5029	-0.7468	-0.9908	-1.2374	0	
	-0.6668	-0.9108	-1.1573	0	0	
	-0.8309	-1.0772	0	0	0	
	-0.9971	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	

And, we can see that the values of $\triangle v(s)$ in Table 3.17 are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion.

If we change λ_n=24, λ_h=12 and keep all the parameters the same as in Table 3.15, this would make the system in a heavier traffic load, the actions and values of $v(\langle s_n, s_h, D\rangle)$ are shown in Table 3.18 and Table 3.19.

Table 3.18: Actions for call arrivals

b=A_n	$s_h \rightarrow$										
	1	1	1	0	0	0	0	0	0	0	0
	1	1	1	0	0	0	0	0	0	0	-1
	1	1	0	0	0	0	0	0	0	-1	-1
	1	0	0	0	0	0	0	0	-1	-1	-1
$s_n \downarrow$	0	0	0	0	0	0	0	-1	-1	-1	-1
	0	0	0	0	0	0	-1	-1	-1	-1	-1
	0	0	0	0	0	-1	-1	-1	-1	-1	-1
	0	0	0	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
b=A_2	$s_h \rightarrow$										
	1	1	1	1	1	1	1	1	1	1	0
	1	1	1	1	1	1	1	1	1	0	-1
	1	1	1	1	1	1	1	1	0	-1	-1
	1	1	1	1	1	1	1	0	-1	-1	-1
	1	1	1	1	1	1	0	-1	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	0	-1	-1	-1	-1	-1
	1	1	1	1	0	-1	-1	-1	-1	-1	-1
	1	1	1	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

As seen from Table 3.18, compared to Table 3.16, the limit to accept the new call is smaller, this is because the system is in a heavier traffic load.

Table 3.19: Total discounted expected reward $v(s)$

b=D	s_h=0	1	2	3	4	5
	6.7284	6.4844	6.2404	5.9964	5.7523	5.5082
	6.5644	6.3204	6.0764	5.8324	5.5883	5.344
	6.4005	6.1565	5.9124	5.6683	5.4241	5.1796
	6.2365	5.9925	5.7484	5.5042	5.2598	5.0144
	6.0725	5.8284	5.5843	5.3399	5.0947	4.8467
$s_n \downarrow$	5.9085	5.6644	5.42	5.175	4.9273	4.6696
	5.7444	5.5001	5.2552	5.0078	4.7506	0
	5.5802	5.3354	5.0882	4.8315	0	0
	5.4023	5.1575	4.9108	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	6	7	8	9	10	
	5.2639	5.0191	4.7732	4.5237	4.263	
	5.0994	4.8536	4.6046	4.3446	0	
	4.934	4.6854	4.426	0	0	
	4.7661	4.5073	0	0	0	
	4.5885	0	0	0	0	
$s_n \downarrow$	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	

We can see that the values of $\triangle v(s)$ in Table 3.19 are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion. Compared to Table 3.17, the values of $v(\langle s_n, s_h, D\rangle)$ is larger, this is because the system is in a heavier traffic load.

If we only change the cost function $f(x, y) = 0.2(x^2 + y^2)$, the actions and values of $v(\langle s_n, s_h, D\rangle)$ are shown in Table 3.20 and Table 3.21.

Table 3.20: Actions for call arrivals

b=A_n	$s_h \rightarrow$									
	1	1	1	1	1	1	1	1	1	1	0
	1	1	1	1	1	1	1	1	1	0	-1
	1	1	1	1	1	1	1	1	0	-1	-1
	1	1	1	1	1	1	1	0	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	1	0	-1	-1	-1	-1
	0	0	0	0	0	0	-1	-1	-1	-1	-1
	0	0	0	0	0	-1	-1	-1	-1	-1	-1
	0	0	0	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
b=A_2	$s_h \rightarrow$										
	1	1	1	1	1	0	0	0	0	0	0
	1	1	1	1	1	0	0	0	0	0	-1
	1	1	1	1	1	0	0	0	0	-1	-1
	1	1	1	1	1	0	0	0	-1	-1	-1
	1	1	1	1	1	0	0	-1	-1	-1	-1
$s_n \downarrow$	1	1	1	1	1	0	-1	-1	-1	-1	-1
	1	1	1	1	0	-1	-1	-1	-1	-1	-1
	1	1	1	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

As seen from Table 3.20, due to the difference in the reward for new calls and handoff calls, the system can accept the new call if there are several handoff calls in the system, but not able to accept more if there are already many new calls in the system.

Table 3.21: Total discounted expected reward $v(s)$

b=D	s_h=0	1	2	3	4	5
	10.3645	10.2916	10.1693	9.9978	9.777	9.5085
	10.31	10.2371	10.1149	9.9433	9.7225	9.4541
	10.2225	10.1496	10.0274	9.8558	9.635	9.3666
	10.102	10.0292	9.9069	9.7353	9.5145	9.2461
	9.9488	9.876	9.7537	9.5821	9.3613	9.0929
$s_n \downarrow$	9.7647	9.6919	9.5696	9.398	9.1773	8.9088
	9.5508	9.478	9.3557	9.1841	8.9621	0
	9.3062	9.2333	9.1109	8.9365	0	0
	9.0207	8.9497	8.8293	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	6	7	8	9	10	
	9.1945	8.8337	8.4255	7.9684	7.4498	
	9.14	8.7792	8.3704	7.9036	0	
	9.0525	8.6914	8.2759	0	0	
	8.9319	8.567	0	0	0	
	8.7775	0	0	0	0	
$s_n \downarrow$	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	
	0	0	0	0	0	

We can see that the values of $\triangle v(s)$ in Table 3.21 are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion.

3.5 Numerical Analysis

In this section we will do some numerical analysis on the simulation results. With these analysis, for the one class model, we figure out the effects of parameters changing on the optimal CAC policies and values of $v(s, a)$. For the two classes models, we compare the NPS and RCS schemes to find their differences on their optimal CAC polices.

3.5.1 One Class Model

In the simulation part we have seen the values of $v(s,a)$ and actions for states under certain parameters setting. Next, we change the value of some parameters to see their effects on the values of $v(s,a)$.

- Rewards R

  Let the reward R changes from 0.5 to 3, and keep all other parameters the same. The two surface plots of $v(s,a), a = a_R, a_A$ and their difference plot are shown in Fig 3-1. In the figure a_R, a_A is displayed as 0,1. So are in the following figures.

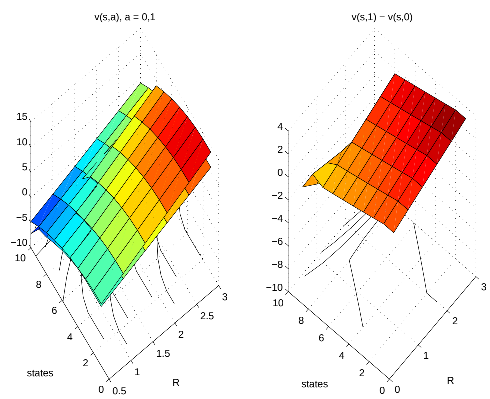

Figure 3-1: $V(s,a)$ as R varies from 0.5 to 3

  It is shown from Fig 3-1 that as R increases, the values of $v(s,a)$ increases. And,

$$v(s,a_A) - v(s,a_R) = R + \triangle v(s,a_R).$$

  As seen from the figure, because $\triangle v(s,a_R)$ is nonincreasing, the difference of $v(s,a_A) - v(s,a_R)$ becomes higher along the reward R axis but drops along the state s axis. And the control limit of states s becomes larger with R increasing.

- Discount Factor α

Let the discount factor α change from 0.1 to 0.9, and all other parameters stay the same. The two surface plots of $v(s,a), a = a_R, a_A$ and their difference plot is shown in Fig 3-2.

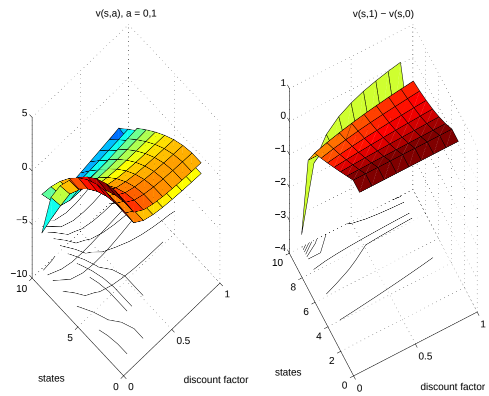

Figure 3-2: $V(s,a)$ as α varies from 0.1 to 0.9

We know for the same interarrival process $A(t)$, $c = \int_0^\infty e^{-\alpha t} dA(t)$ becomes smaller as α increases, so the values for states $s \leq m$ which is $v(s) = \frac{R}{1-c}$ becomes smaller. And as the discount factor α becomes larger, the effect of future reward and waiting cost are weakened, so the values of $\triangle V(s,a)$ are smaller. This can be seen from Fig 3-2.

- Traffic Load $\frac{\lambda}{m\mu}$

  Also, let the Arrival Rate λ changes from 0.5 to 4, and all other parameters stay the same. The two surface plots of $v(s,a)$ are shown in Fig 3-3.

  Since $A(t)$ is an $Erlang-k$ distribution with $k = 5$, as the arrival rate increases, $c = \int_0^\infty e^{-\alpha t} dA(t)$ becomes larger, so the values for states $s \leq m$ which is $v(s) = \frac{R}{1-c}$ becomes larger. As shown in Figure 3-3, if the traffic load is heavier, the control limit is smaller. This means the system starts rejecting calls at a lower state.

The above analysis has revealed that:

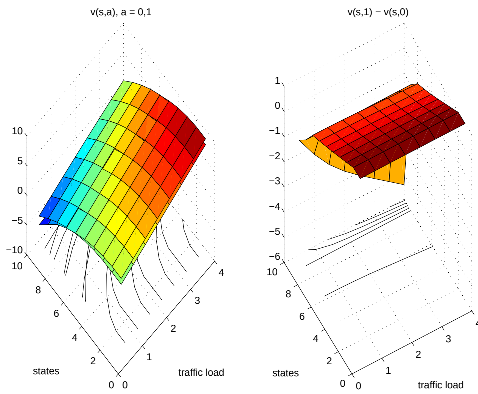

Figure 3-3: $V(s,a)$ as $\frac{\lambda}{m\mu}$ varies from 1/8 to 1

1. Random initial policy leads to the same control limit policy. This proves the correctness of our search algorithm.

2. When R is fixed, increasing channel or buffer capacities does not change the optimal policy.

3. It is better to increase the number of channels if the traffic load becomes heavier.

For simple computation, we only did the simulation of $m = 2$ and $C = 10$. But by using our algorithm, these could easily be extended to large numbers to find the optimal control limit. Also, based on the values of $V(s,a)$ we can design our system to meet some special requirements, like the number of channels that can bring us the best reward when all other parameters are known.

3.5.2 Comparisons on NPS and RCS

Compare the two SMDP models for the RCS and NPS scheme, these two models have the same action sets, the same probability transitions, the same rewards, the same decision epochs, and only a little difference in the **state space**.

In this section, we investigate the difference of the optimal policy for the RCS and NPS schemes. To make the comparison, we focused on the values of total expected discounted reward $v(s)$ and the actions for each state. To further compare the performance of these two schemes on the rewards, a new concept of total expected discounted system reward is proposed. Also, we show the effect on the optimal policy for the NPS scheme if the cost function does not follow the property requirement in Theorem 3.7.

Actions for calls

In Table 3.22, '—' means to admit a handoff call in such state, | is for admitting a new call and '+' is for both calls. '0' are those states in which the system is full. '-1' denotes those states that do not exist in the system. From Table 3.22, we see that in both schemes, the system admits the handoff calls whenever there is a free channel, but rejects the new call if there are already several calls in the system. This decision rule is due to the difference in the rewards of new calls and handoff calls. The optimal policy is a control limit policy, which is consistent with Theorem 3.7.

And, for states with $s_h = 0$, the RCS scheme rejects the new call at $s_n = 7$, which is a little earlier than the NPS scheme which rejects the new call at $s_n = 8$. Because there are 2 channels reserved for the handoff calls, the RCS scheme rejects the new call earlier to save the space for handoff calls, which gives the handoff calls some more priority.

Total expected discounted reward $v(s)$

Using the parameters setting in Table 3.15, the difference of values of total expected discounted reward $v(s)$ of states $b = D$ between the RCS and NPS schemes are plotted in Fig. 3-4.

From Fig. 3-4, we can see that although there are difference in the state space for the RCS and NPS schemes, there is not much difference in the values of total expected discounted reward $v(s)$ for the common states which exist in both two schemes (except the states existing in NPS but not in RCS). This is the same as shown in Fig. 3-5, where we increased $\lambda_n = 24$, and $\lambda_h = 12$ to make the system in a heavy traffic load.

Table 3.22: Actions for call arrivals

RCS	$s_h \rightarrow$									
	+	+	+	+	+	+	+	+	—	—	0
	+	+	+	+	+	+	+	—	—	0	-1
	+	+	+	+	+	+	—	—	0	-1	-1
	+	+	+	+	+	—	—	0	-1	-1	-1
	+	+	+	+	—	—	0	-1	-1	-1	-1
$s_n \downarrow$	+	+	+	—	—	0	-1	-1	-1	-1	-1
	+	+	—	—	0	-1	-1	-1	-1	-1	-1
	—	—	—	0	-1	-1	-1	-1	-1	-1	-1
	0	0	0	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
NPS	$s_h \rightarrow$										
	+	+	+	+	+	+	+	+	—	—	0
	+	+	+	+	+	+	+	—	—	0	-1
	+	+	+	+	+	+	—	—	0	-1	-1
	+	+	+	+	+	—	—	0	-1	-1	-1
	+	+	+	+	—	—	0	-1	-1	-1	-1
$s_n \downarrow$	+	+	+	—	—	0	-1	-1	-1	-1	-1
	+	+	—	—	0	-1	-1	-1	-1	-1	-1
	+	—	—	0	-1	-1	-1	-1	-1	-1	-1
	—	—	0	-1	-1	-1	-1	-1	-1	-1	-1
	—	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Also, we can see that the values of $\triangle v(s)$ are nonincreasing in both $s_n \downarrow$ and $s_h \rightarrow$ directions, which fits our theoretical conclusion of Theorem 3.7.

Different holding cost function

In previous sections of this paper we set the cost function as $f(s_n, s_h) = (s_n + s_h)$, which fits the requirements of Theorem 3.7. Now we make some changes to the cost functions and see what will happen to the optimal policy? For example, let us set the cost function as

$$f(s_n, s_h) = \begin{cases} s_n + s_h, & s_h \leq 3, \\ 2s_n + s_h, & s_h > 3. \end{cases}$$

So $f(s_n, s_h)$ is convex nondecreasing on s_n but not on s_h, there is a break point at

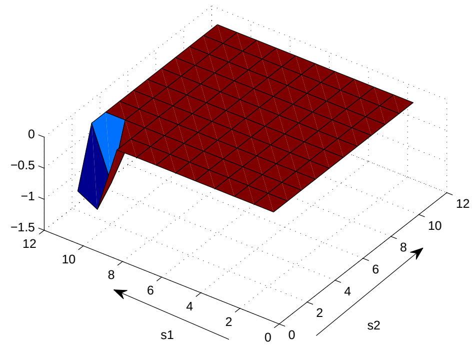

Figure 3-4: Difference on $V(s, a)$

$s_h = 3$. All the other parameters remained the same. The decision rule for the NPS scheme are shown in Table 3.23.

As seen from Table 3.23, the optimal policy is not a control limit policy. The new calls are rejected if there are more than 3 handoff calls in the system, because its holding cost rate is doubled. For the handoff calls, the actions are to accept when $s_h \leq 2$. However from $s_h \geq 3$, because of the doubled holding cost rate for the new calls, the system rejects the handoff call when there are a few new calls in the system. As the number of handoff calls increases, which brings more reward to cover the holding cost, the system accepts the handoff calls even when there are several new calls in the system.

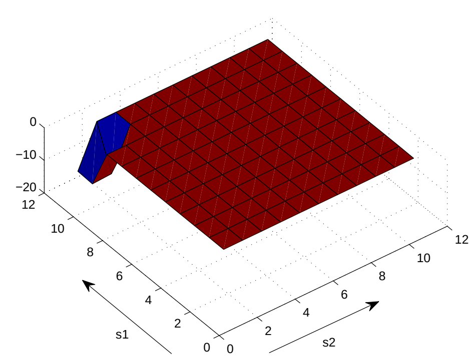

Figure 3-5: Difference on $V(s, a)$, $\lambda_n = 24$, $\lambda_h = 12$

Table 3.23: Actions for call arrivals

NPS	s_h $\rightarrow$										
	+	+	+	—	—	—	—	—	—	—	0
	+	+	$\vert$	—	—	—	—	—	—	0	-1
	+	+	$\vert$	0	—	—	—	—	0	-1	-1
	+	+	$\vert$	0	0	—	—	0	-1	-1	-1
	+	+	$\vert$	0	0	—	0	-1	-1	-1	-1
$s_n \downarrow$	+	+	$\vert$	0	0	0	-1	-1	-1	-1	-1
	+	+	$\vert$	0	0	-1	-1	-1	-1	-1	-1
	+	+	$\vert$	0	-1	-1	-1	-1	-1	-1	-1
	+	+	0	-1	-1	-1	-1	-1	-1	-1	-1
	+	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

We examined the difference between the optimal policies of two different resource allocation schemes: Non-Priority Scheme (NPS) and Reserved Channel Scheme (RCS) using SMDP models. We assumed there is some reward for accepting each call and holding cost rate per unit time for the calls in the system. We find that if the cost functions $f(s_n, s_h)$ have some special properties, the optimal policies for both schemes are control limit policies. In addition we investigate the difference of total expected discounted reward $v(s)$ and actions decision rules generated by the optimal policies of two schemes through numerical analysis. From our results we conclude that the two schemes, NPS and RCS, demonstrate an almost equal performance in achieving maximum rewards in a wireless network.

Chapter 4

Modeling, Simulation and Analysis of Multimedia Traffic In a Cell

As the first two generations of wireless networks were mainly designed for providing voice (and low-rate data) services, the third generation (3G) wireless networks (such as cdma2000) is to provide high rate integrated services, with data rates up to 2M bps for a single user. Applications such as Internet access and video conference can thus be supported over wireless. As we know, the next-generation cellular systems will support multimedia services such as voice, data, facsimiles and video, effective traffic control schemes for multimedia traffic with various characteristics are needed. The wide range of services is characterized by different bit rates and statistical behavior of multimedia traffic sources and by different quality of service (QoS) requirements.

4.1 Introduction

In the last chapter we studied on a single cell with homogeneous traffic calls. In this chapter we consider a single cell in a multimedia wireless network. Since the handoff calls can be looked as a special class of calls, we do not mention them out in the model definitions. In the first section we study on the NPS schemes, here the idea of non priority is extended to all the classes of calls in the system, which means any call from any class is treated equally in the system. In the second section we study on the RCS schemes, where some channels are reserved for the calls of some classes, e.g the calls in the handoff classes.

A wireless multimedia network has to be able to support multiple classes of traffic (voice, video, images, etc) with different Quality of Service (QoS) requirements, i.e., different number of channels needed, holding time of the connection and cell residence time. Without losing generosity, suppose there are U classes of multimedia calls (telephone, video, etc., but for convenience we shall call all of them as *calls*) in the wireless network.

4.2 Non Priority Scheme

For every class of calls, assuming the new traffic origination rates are uniformly distributed over the mobile service area. Any incoming call to the cell can get a service if there are enough free channels in the system.

4.2.1 Model Formulation

The other assumptions and notations are as follows:

1. The required bandwidth of class u calls $u = 1, \ldots, U$ is b_u.

2. Class u calls are generated according to a Poisson process with rates λ_u, $u = (1, \ldots, U)$.

3. The total bandwidth in the network is C. The service time of class-u calls follows negative exponential distribution with rate μ_u.

4. Accepting each incoming class-u call would contribute R_u units of reward to the system. For the calls exceeding the reserved channels, the system incurs a holding cost rate $f()$ per unit time.

Each class calls need some channels in service. Here the capacity C mainly refers to the bandwidth of a network, described in the number of channels. Based on these assumptions, we can build the SMDP model for the system as follows:

1. Let the state variable consists of number of calls of different classes in the system, the status of calls leaving, arriving to the system, so $S = \{0, 1, \ldots, \lfloor \frac{C}{b_1} \rfloor\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C}{b_U} \rfloor\} \times \{D, A_1, \ldots, A_U\}$, where $\lfloor x \rfloor$ denotes the nearest integer to x. And in the third set D means a departure from the system, while A_u means an arrival of a class-u call. And, $\sum_{u=1}^{U} n_u b_u \leq C$. In general a state could be written as $\langle \mathbf{n}, b \rangle$, where $\mathbf{n} = (n_1, \ldots, n_U)$ are the numbers of calls in each class and b stands for the last call event, and $b \in \{D, A_1, \ldots, A_U\}$.

2. In states $\langle \mathbf{n}, D \rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D \rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_u \rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_u \rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s, a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b \rangle$ and action a, let $\beta_0 = \sum_{u=1}^{N} (\lambda_u + n_u \mu_u)$, so $\beta(s, a)$ can be written as

$$\beta(s,a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + \mu_u, & a = a_A, b = A_u. \end{cases}$$

4. Let $q(j|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle n, b \rangle$, $b = D, a = a_C$ and $b = \{A_u, u = (1, 2, \cdots, U)\}, a = a_R$, let $\mathbf{n}_u = \{n_1, n_2, \ldots, \max(n_u - 1, 0), \ldots, n_N\}$ and $\mathbf{n}^u = \{n_1, n_2, \ldots, n_u + 1, \ldots, n_N\}$, the state transition probabilities are

$$q(j|s,a) = \begin{cases} \lambda_u/\beta_0, & j = \langle \mathbf{n}, A_u \rangle, \\ n_u\mu_u/\beta_0, & j = \langle \mathbf{n}_u, D \rangle, \end{cases}$$

And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_u, u = (1, 2, \cdots, U), a = a_A$, since admitting an incoming call migrates the system state immediately, we have

$$q(j|s, a_A) = q(j|\langle \mathbf{n}^u, D \rangle, a_R), b = A_u$$

5. Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a)E_s^a \left\{ \int_0^t e^{-\alpha\tau} d\tau \right\} \\ &= k(s,a) + c(s,a)E_s^a \left\{ [1 - e^{-\alpha t}]/\alpha \right\} \\ &= k(s,a) + \frac{c(s,a)}{\alpha + \beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \\ R_u, & a = a_A, b = A_u. \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(\mathbf{n}), & a = a_C, a = a_R, \\ -f(\mathbf{n}^u), & a = a_A, b = A_u. \end{cases}$$

$f(\mathbf{n}) = 0$ if $n_i \leq k_i, i = 1, \ldots, U$, and

$f(\mathbf{n}) = f(n_1, \ldots, n_i, \ldots, n_U)) = f(n_1, \ldots, 0, \ldots, n_U))$ if $n_i \leq k_i$.

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(j|s) = q(j|s, d(s))$, $r_d(s) = r(s, d(s))$ and $\beta_d(s) = \beta(s, d(s))$, the expected infinite-horizon discounted reward with discount factor α is

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{j \in S} q_d(j|s) v_\alpha^{d^\infty}(j). \tag{4.1}$$

From equation (4.1), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, this process fits the condition of $[1 - q(s|s,a)]\beta(s,a) \leq c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of this process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = \lambda_1 + \lambda_2 + C * \max(\mu_1, \mu_2)$, we have

$$\tilde{q}(j|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & j = s, \\ \frac{q(j|s,a)\beta(s,a)}{c}, & j \neq s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha + \beta(s,a)}{\alpha + c}.$$

For each stationary deterministic d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{4.2}$$

From equation (4.1) and (4.2), the optimality equation of $v(s)$ for maximum $v_\alpha^{d^\infty}(s)$ would have the form of

$$v(s) = \max_{a \in A_s} \left\{ \tilde{r}(s,a) + \lambda \sum_{j \in S} \tilde{q}(j|s,a)v(j) \right\}, \tag{4.3}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (4.3), for states with $b = D$, since there is only one action a_c for these states, we have

$$\begin{aligned} v(\langle \mathbf{n}, D \rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) &+ \sum_{u=1}^{U} n_u \mu_u v(\langle \mathbf{n}_u, D \rangle) \\ + \sum_{u=1}^{U} \lambda_u v(\langle \mathbf{n}, A_u \rangle) &+ (c - \beta_0) v(\langle \mathbf{n}, D \rangle)]. \end{aligned} \tag{4.4}$$

4.2.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_u$,

$$v(\langle \mathbf{n}, A_u \rangle) = \max\left[v(\langle \mathbf{n}, D \rangle), R_u + v(\langle \mathbf{n}^u, D \rangle)\right].$$

Theorem 4.1 *If $v(\langle \boldsymbol{n}, D \rangle)$ is concave nonincreasing on n_u, $u = 1, \ldots, U$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_u(\langle \mathbf{n}, D \rangle) = v(\langle \mathbf{n}^u, D \rangle) - v(\langle \mathbf{n}, D \rangle)$, so $\triangle v_u(\langle \mathbf{n}, D \rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_u, we have the decision rule

$$d(n, A_u) = \begin{cases} a_A, & \triangle v_u(\langle \mathbf{n}, D \rangle) > -R_u, \\ a_R, & \triangle v_u(\langle \mathbf{n}, D \rangle) \leq -R_u. \end{cases}$$

So, if $d(\mathbf{n}, A_u) = a_R$, we have $d(\mathbf{n}^u, A_u) = a_R$, and so on as u goes through $1, \ldots, U$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 4.2 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_u, u = 1, \ldots, U$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{ux}) \geq f(\boldsymbol{n}^u) + f(\boldsymbol{n}^x),$$

where $\boldsymbol{n}^{ux} = \{n_1, \ldots, n_u + 1, \ldots, n_x + 1, \ldots, n_N\}, u, x \in 1, \ldots, U$, then (1): $v(\langle \boldsymbol{n}, D \rangle)$ is concave nonincreasing on $n_u, u = 1, \ldots, U$, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (4.4) to obtain $v^1(\langle \mathbf{n}, D \rangle) = \frac{-f(\mathbf{n})}{\alpha + c}$.

2. Set i=1. We have $v^i(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^i(\langle \mathbf{n}, A_u\rangle)$ is concave nonincreasing on $n_u, u = 1, \ldots, U$. And we have,

$$v^i(\langle \mathbf{n}, b\rangle) + v^i(\langle \mathbf{n}^{ux}, b\rangle) \leq v^i(\langle \mathbf{n}^u, b\rangle) + v^i(\langle \mathbf{n}^x, b\rangle),$$

for $b = A_u, u, x \in 1, \ldots, U$.

3. Substituting these v^i back into equation (4.4), we get

$$\begin{aligned} v^{i+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[&-f(\mathbf{n}) + \sum_{u=1}^{U} \lambda_u v^i(\langle \mathbf{n}, A_u\rangle) \\ &+ \sum_{u=1}^{U} n_u \mu_u v^i(\langle \mathbf{n}_u, D\rangle) + (c - \beta_0) v(\langle \mathbf{n}, D\rangle)]. \end{aligned}$$

so we have for $u = 1, \ldots, U$, let $\mathbf{n}_x^u = \{n_1, \ldots, max(n_x - 1, 0), \ldots, n_u + 1, \ldots, n_N\}$

$$\begin{aligned} &v^{i+1}(\langle \mathbf{n}^u, D\rangle) - v^{i+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[\\ &\quad \sum_{w=1}^{U} \lambda_w (v^i(\langle \mathbf{n}^u, A_w\rangle) - v^i(\langle \mathbf{n}, A_w\rangle)) + (f(\mathbf{n}) - f(\mathbf{n}^u)) \\ &\quad + \sum_{w=1}^{U} n_w \mu_w (v^i(\langle \mathbf{n}_w^u, D\rangle) - v^i(\langle \mathbf{n}_w, D\rangle)) \\ &\quad + (c - \beta_u)(v^i(\langle \mathbf{n}^u, D\rangle) - v^i(\langle \mathbf{n}, D\rangle))], \end{aligned}$$

where $\beta_u = \beta_0 + \mu_u$. Since $v^i(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_u, the combination of these functions $v^{i+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_u. Also, using equation (4.4),

$$\begin{aligned} &v^{i+1}(\langle \mathbf{n}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{ux}, D\rangle) \\ &- v^{i+1}(\langle \mathbf{n}^u, D\rangle) - v^{i+1}(\langle \mathbf{n}^x, D\rangle) \end{aligned}$$

could be divided into several different parts of $v^i(\langle \mathbf{n}, b\rangle) + v^i(\langle \mathbf{n}^{ux}, b\rangle) - v^i(\langle \mathbf{n}^u, b\rangle) -$

$v^i(\langle \mathbf{n}^x, b\rangle), b = D, A_u,\ u = 1, \ldots, U$ which are all ≤ 0. So

$$\begin{aligned} &v^{i+1}(\langle \mathbf{n}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{ux}, D\rangle) \\ \leq\ &v^{i+1}(\langle \mathbf{n}^{u}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{x}, D\rangle). \end{aligned}$$

4. Set i=i+1, go back to step 2.

5. As the iteration continues, $v^i(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^i(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 4.1, the optimal policy is a control limit policy. The verification is over.

4.2.3 Simulation Results

Without losing generality and for computation simplification, suppose there are 2 classes of calls in this wireless mobile network, let the total channels (capacity) be $C = 6$, the cost function be $f(x, y) = x^2 + y^2$, the discount factor $\alpha = 0.1$, and we set the other parameters for analysis as in Table 4.1 to study the performance of the RCS schemes.

Table 4.1: Parameters

Class	b	λ	μ	R
1	1	4	2	2
2	2	2	3	1.5

In Table 4.1, b is the bandwidth or the number of channels, λ is the arrival rate of calls, μ is the call service rate, R is the reward for accepting calls.

We get the actions and $v(s)$ values for states in the Table 4.2 and Table 4.3.

Table 4.2: Actions for calls

b=A_1	$n_2 \rightarrow$			
$n_1 \downarrow$	1	1	1	Φ
	1	1	1	-1
	1	0	Φ	-1
	0	0	-1	-1
	0	Φ	-1	-1
	0	-1	-1	-1
	Φ	-1	-1	-1
b=A_2	$n_2 \rightarrow$			
$n_1 \downarrow$	1	1	0	Φ
	1	1	Φ	-1
	1	1	Φ	-1
	1	Φ	-1	-1
	1	Φ	-1	-1
	Φ	-1	-1	-1
	Φ	-1	-1	-1

It is seen that in Table 4.2, '1' means to accept the call, '0' means to reject the call on such state, Φ are those states in which the system is full or the system can not take this class call anymore, and '-1' means that this state does not exist in the state space. It can be seen in Table 4.2 that there are limits to accepts the calls, which fits our theoretical conclusion.

We can see that the values of $\triangle v(s)$ in Table 4.3 are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion.

Table 4.3: $v(s)$ for states

b=D	$n_2 \rightarrow$			
$n_1 \downarrow$	48.6426	48.0678	47.1132	45.6065
	47.3961	46.8158	45.8591	0
	45.7479	45.1532	44.1936	0
	43.7908	43.1090	0	0
	41.4670	40.7453	0	0
	38.5812	0	0	0
	35.2871	0	0	0
b=A_1	$n_2 \rightarrow$			
$n_1 \downarrow$	49.3961	48.8158	47.8591	45.6065
	47.7479	47.1532	46.1936	0
	45.7908	45.1532	44.1936	0
	43.7908	43.1090	0	0
	41.4670	40.7453	0	0
	38.5812	0	0	0
	35.2871	0	0	0
b=A_2	$n_2 \rightarrow$			
$n_1 \downarrow$	49.5678	48.6132	47.1132	45.6065
	48.3158	47.3591	45.8591	0
	46.6532	45.6936	44.1936	0
	44.6090	43.1090	0	0
	42.2453	40.7453	0	0
	38.5812	0	0	0
	35.2871	0	0	0

If we increase the channel capacity to C=10 and keep all the parameters the same as in Table 4.1,the actions are shown in Table 4.4.

As seen from Table 4.4, compared to Table 4.2, the limits to accept the new calls are not changed much. This shows that simply increasing channel capacity is not a good way to increase the thresholds for accepting calls.

Table 4.4: Actions for call arrivals

b=A_1	$n_2 \rightarrow$					
	1	1	1	1	1	0
	1	1	1	1	1	-1
	1	1	1	1	0	-1
	0	0	0	0	-1	-1
	0	0	0	0	-1	-1
$n_1 \downarrow$	0	0	0	-1	-1	-1
	0	0	0	-1	-1	-1
	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1
b=A_2	$n_2 \rightarrow$					
	1	1	1	1	0	0
	1	1	1	1	0	-1
	1	1	1	1	0	-1
	1	1	1	0	-1	-1
	1	1	1	0	-1	-1
$n_1 \downarrow$	1	1	0	-1	-1	-1
	1	1	0	-1	-1	-1
	1	0	-1	-1	-1	-1
	1	0	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1

4.3 Reserved Channel Scheme

We take handoff calls as a special class of class in the multimedia network. Since it is always more desirable not to drop a handoff call than a new call, some classes of calls are more important or may bring more benefit to the provider, we reserve some channels for these classes of calls and still use the name of 'Reserved Channel Scheme'.

4.3.1 Model Formulation

The other assumptions and notations are as follows:

1. The required bandwidth of class u calls $u = 1, \ldots, U$ is b_u.

2. To be fair to each class of calls in the system, there are $K_u b_u (K_u \geq 0)$ channels reserved for class u calls.

3. Class u calls are generated according to a Poisson process with rates λ_u, $u = (1, \ldots, U)$.

4. The total bandwidth in the network is C. The service time of class-u calls follows negative exponential distribution with rate μ_u.

5. Accepting each incoming class-u call would contribute R_u units of reward to the system. For the calls exceeding the reserved channels, the system incurs a holding cost rate $f()$ per unit time.

Based on these assumptions, we can build the SMDP model for the system as follows:

1. Let the state variable consists of number of calls of different classes in the system, the status of calls leaving, arriving to the system. Let $K = \sum_{u=1}^{U} K_u b_u$, which is the total channels reserved for all classes, then $S = \{0, 1, \ldots, \lfloor \frac{C-K}{b_1} \rfloor + K_1\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C-K}{b_U} \rfloor + K_U\} \times \{D, A_1, \ldots, A_U\}$, where $\lfloor x \rfloor$ denotes the nearest integer to x. And in the third set D means a departure from the system, while A_u means an arrival of a class-u call. And, $\sum_{u=1}^{U} n_u b_u \leq C$. In general a state could be written as $\langle \mathbf{n}, b \rangle$, where $\mathbf{n} = (n_1, \ldots, n_U)$ are the numbers of calls in each class and b stands for the last call event, and $b \in \{D, A_1, \ldots, A_U\}$.

2. In states $\langle \mathbf{n}, D \rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D \rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_u \rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_u \rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s,a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b \rangle$ and action a, let $\beta_0 = \sum_{u=1}^{N}(\lambda_u + n_u\mu_u)$, so $\beta(s,a)$ can be written as

$$\beta(s,a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + \mu_u, & a = a_A, b = A_u. \end{cases}$$

4. Let $q(j|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle n, b \rangle$, $b = D, a = a_C$ and $b = \{A_u, u = (1, 2, \cdots, U)\}, a = a_R$, let $\mathbf{n}_u = \{n_1, n_2, \ldots, \max(n_u - 1, 0), \ldots, n_N\}$ and $\mathbf{n}^u = \{n_1, n_2, \ldots, n_u + 1, \ldots, n_N\}$, the state transition probabilities are

$$q(j|s,a) = \begin{cases} \lambda_u/\beta_0, & j = \langle \mathbf{n}, A_u \rangle, \\ n_u\mu_u/\beta_0, & j = \langle \mathbf{n}_u, D \rangle, \end{cases}$$

And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_u, u = (1, 2, \cdots, U), a = a_A$, since admitting an incoming call migrates the system state immediately, we have

$$q(j|s, a_A) = q(j|\langle \mathbf{n}^u, D \rangle, a_R), b = A_u$$

5. Because the system state does not change between decision epochs, the expected

discounted reward between epochs satisfies

$$
\begin{aligned}
r(s,a) &= k(s,a) + c(s,a) E_s^a \left\{ \int_0^t e^{-\alpha\tau} d\tau \right\} \\
&= k(s,a) + c(s,a) E_s^a \left\{ [1 - e^{-\alpha t}]/\alpha \right\} \\
&= k(s,a) + \frac{c(s,a)}{\alpha + \beta(s,a)},
\end{aligned}
$$

where

$$
k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \\ R_u, & a = a_A, b = A_u. \end{cases}
$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$
c(s,a) = \begin{cases} -f(\mathbf{n}), & a = a_C, a = a_R, \\ -f(\mathbf{n}^u), & a = a_A, b = A_u. \end{cases}
$$

$f(\mathbf{n}) = 0$ if $n_i \le k_i, i = 1, \ldots, U$, and

$f(\mathbf{n}) = f(n_1, \ldots, n_i, \ldots, n_U)) = f(n_1, \ldots, 0, \ldots, n_U))$ if $n_i \le k_i$.

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(j|s) = q(j|s, d(s))$, $r_d(s) = r(s, d(s))$ and $\beta_d(s) = \beta(s, d(s))$, the expected infinite-horizon discounted reward with discount factor α is

$$
v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{j \in S} q_d(j|s) v_\alpha^{d^\infty}(j). \tag{4.5}
$$

From equation (4.5), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, this process fits the condition of $[1 - q(s|s,a)]\beta(s,a) \leq c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of this process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = \lambda_1 + \lambda_2 + C * \max(\mu_1, \mu_2)$, we have

$$\tilde{q}(j|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & j = s, \\ \frac{q(j|s,a)\beta(s,a)}{c}, & j \neq s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha + \beta(s,a)}{\alpha + c}.$$

For each stationary deterministic d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{4.6}$$

From equation (4.5) and (4.6), the optimality equation of $v(s)$ for maximum $v_\alpha^{d^\infty}(s)$ would have the form of

$$v(s) = \max_{a \in A_s} \left\{ \tilde{r}(s,a) + \lambda \sum_{j \in S} \tilde{q}(j|s,a)v(j) \right\}, \tag{4.7}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (4.7), for states with $b = D$, since there is only one action a_c for these states, we have

$$v(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{u=1}^{U} n_u \mu_u v(\langle \mathbf{n}_u, D\rangle)$$
$$+ \sum_{u=1}^{U} \lambda_u v(\langle \mathbf{n}, A_u\rangle) + (c - \beta_0) v(\langle \mathbf{n}, D\rangle)]. \tag{4.8}$$

4.3.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_u$,

$$v(\langle \mathbf{n}, A_u\rangle) = \max\left[v(\langle \mathbf{n}, D\rangle), R_u + v(\langle \mathbf{n}^u, D\rangle)\right].$$

Theorem 4.3 *If $v(\langle \boldsymbol{n}, D\rangle)$ is concave nonincreasing on n_u, $u = 1, \ldots, U$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_u(\langle \mathbf{n}, D\rangle) = v(\langle \mathbf{n}^u, D\rangle) - v(\langle \mathbf{n}, D\rangle)$, so $\triangle v_u(\langle \mathbf{n}, D\rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_u, we have the decision rule

$$d(n) = \begin{cases} a_A, & \triangle v_u(\langle \mathbf{n}, D\rangle) > -R_u, \\ a_R, & \triangle v_u(\langle \mathbf{n}, D\rangle) \leq -R_u. \end{cases}$$

So, if $d(\mathbf{n}) = a_R$, we have $d(\mathbf{n}^u) = a_R$, and so on as u goes through $1, \ldots, U$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 4.4 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_u, u = 1, \ldots, U$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{ux}) \geq f(\boldsymbol{n}^u) + f(\boldsymbol{n}^x),$$

where $\boldsymbol{n}^{ux} = \{n_1, \ldots, n_u + 1, \ldots, n_x + 1, \ldots, n_N\}, u, x \in 1, \ldots, U$, *then (1):* $v(\langle \boldsymbol{n}, D\rangle)$ *is concave nonincreasing on* $n_u, u = 1, \ldots, U$, *and (2): the optimal policy is a control limit policy.*

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (4.8) to obtain $v^1(\langle \mathbf{n}, D\rangle) = \frac{-f(\mathbf{n})}{\alpha+c}$.

2. Set i=1. We have $v^i(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^i(\langle \mathbf{n}, A_u\rangle)$ is concave nonincreasing on $n_u, u = 1, \ldots, U$. And we have,

$$v^i(\langle \mathbf{n}, b\rangle) + v^i(\langle \mathbf{n}^{ux}, b\rangle) \leq v^i(\langle \mathbf{n}^u, b\rangle) + v^i(\langle \mathbf{n}^x, b\rangle),$$

   for $b = A_u, u, x \in 1, \ldots, U$.

3. Substituting these v^i back into equation (4.8), we get

$$\begin{aligned} v^{i+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{u=1}^{U} \lambda_u v^i(\langle \mathbf{n}, A_u\rangle) \\ + \sum_{u=1}^{U} n_u \mu_u v^i(\langle \mathbf{n}_u, D\rangle) + (c - \beta_0) v(\langle \mathbf{n}, D\rangle)]. \end{aligned}$$

   so we have for $u = 1, \ldots, U$, let $\mathbf{n}_x^u = \{n_1, \ldots, max(n_x - 1, 0), \ldots, n_u + 1, \ldots, n_N\}$

$$\begin{aligned} v^{i+1}(\langle \mathbf{n}^u, D\rangle) - v^{i+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[\\ \sum_{w=1}^{U} \lambda_w (v^i(\langle \mathbf{n}^u, A_w\rangle) - v^i(\langle \mathbf{n}, A_w\rangle)) + (f(\mathbf{n}) - f(\mathbf{n}^u)) \\ + \sum_{w=1}^{U} n_w \mu_w (v^i(\langle \mathbf{n}_w^u, D\rangle) - v^i(\langle \mathbf{n}_w, D\rangle)) \\ +(c - \beta_u)(v^i(\langle \mathbf{n}^u, D\rangle) - v^i(\langle \mathbf{n}, D\rangle))], \end{aligned}$$

   where $\beta_u = \beta_0 + \mu_u$. Since $v^i(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_u, the combination of these functions $v^{i+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_u.

Also, using equation (4.8),

$$\begin{aligned} & v^{i+1}(\langle \mathbf{n}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{ux}, D\rangle) \\ & -v^{i+1}(\langle \mathbf{n}^{u}, D\rangle) - v^{i+1}(\langle \mathbf{n}^{x}, D\rangle) \end{aligned}$$

could be divided into several different parts of $v^i(\langle \mathbf{n}, b\rangle)+v^i(\langle \mathbf{n}^{ux}, b\rangle)-v^i(\langle \mathbf{n}^{u}, b\rangle)-v^i(\langle \mathbf{n}^{x}, b\rangle), b = D, A_u,\ u = 1, \ldots, U$ which are all ≤ 0. So

$$\begin{aligned} & v^{i+1}(\langle \mathbf{n}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{ux}, D\rangle) \\ \leq\ & v^{i+1}(\langle \mathbf{n}^{u}, D\rangle) + v^{i+1}(\langle \mathbf{n}^{x}, D\rangle). \end{aligned}$$

4. Set i=i+1, go back to step 2.

5. As the iteration continues, $v^i(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^i(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 4.3, the optimal policy is a control limit policy. The verification is over.

It is seen from Theorem 4.4 that no matter how many channels are reserved, the optimal policy is always a control limit policy if the holding cost functions are convex and nondecreasing.

4.3.3 Simulation Results

Without losing generality and for computation simplification, suppose there are 2 classes of calls in this wireless mobile network, let the total channels (capacity) be $C = 6$, the cost function be $f(x, y) = x^2 + y^2$, the discount factor $\alpha = 0.1$, and we set the other parameters for analysis as in Table 4.5 to study the performance of the

RCS schemes. We get the actions and $v(s)$ values for states in the Table 4.6 and Table 4.7.

Table 4.5: Parameters

Class	b	λ	μ	R	K
1	1	4	2	2	0
2	2	2	3	1.5	1

In Table 4.5, b is the bandwidth or the number of channels, λ is the arrival rate of calls, μ is the call service rate, R is the reward for accepting calls, and K is the number of reserved channels. Table 4.6 shows the actions for calls. In Table 4.6,

Table 4.6: Actions for calls

b=A_1	$n_2 \rightarrow$			
$n_1 \downarrow$	1	1	Φ	Φ
	1	1	Φ	-1
	1	Φ	Φ	-1
	0	Φ	-1	-1
	Φ	Φ	-1	-1
	Φ	-1	-1	-1
	Φ	-1	-1	-1
b=A_2	$n_2 \rightarrow$			
$n_1 \downarrow$	1	1	0	Φ
	1	1	Φ	-1
	1	1	Φ	-1
	1	Φ	-1	-1
	1	Φ	-1	-1
	Φ	-1	-1	-1
	Φ	-1	-1	-1

'1' means to admit a new call and '0' is to reject. Φ are those states in which the system is full or the system can not take this class call anymore, and '-1' denotes those states which does not exist in the system. Due to the difference on bandwidth and reserved number, these states are not the same for class 1 and class 2 arrivals. For example, $(0, 2)$ is a state full for class 1 but not for class 2. '-1' denotes those states that do not exist in the state space. As seen from Table 4.6, we can see that when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the calls, so the optimal policy is a control limit policy, which is consistent with our theorem conclusion.

Table 4.7: $v(s)$ for states

b=D	$n_2 \rightarrow$			
$n_1 \downarrow$	50.3238	49.7757	48.3040	46.7842
	49.1059	48.5949	47.3059	0
	47.4801	46.9807	45.8523	0
	45.5410	45.0008	0	0
	43.2359	42.6769	0	0
	0	0	0	0
	0	0	0	0
b=A_1	$n_2 \rightarrow$			
$n_1 \downarrow$	51.1059	50.5949	48.3040	46.7842
	49.4801	48.9807	47.3059	0
	47.5410	46.9807	45.8523	0
	45.5410	45.0008	0	0
	43.2359	42.6769	0	0
	0	0	0	0
	0	0	0	0
b=A_2	$n_2 \rightarrow$			
$n_1 \downarrow$	51.2757	49.8040	48.3040	46.7842
	50.0949	48.8059	47.3059	0
	48.4807	47.3523	45.8523	0
	46.5008	45.0008	0	0
	44.1769	42.6769	0	0
	0	0	0	0
	0	0	0	0

In Table 4.7, '0' denotes the states that do not exist in the state space. From the Table 4.7 we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion.

4.4 Numerical Analysis

As seen from Table 4.1 and Table 4.5, the only difference is the reserved numbers of channels for different classes. From Table 4.2, when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the calls. Compared with Table 4.6, we see that with the different numbers of reserved channels, the state spaces and actions for states of each SMDP models are different. Also, from the Table 4.7 we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions. Due to difference in the reserved channels, the values in Table 4.3 and Table 4.7 are different. Although there are differences in the actions and $v(s)$ values, the optimal policy obtained is still a control limit policy, which is consistent with our theorem conclusion.

Next, using the parameters setting for reserved channel schemes, without changing other parameters, if we set the cost function to $f(x, y) = x^2 + y^2 - xy$, which does not follow the requirement of Theorem, the actions for states are shown in Table 4.8. As shown from Table 4.8, the optimal policy is not a control limit policy.

Table 4.8: Actions for calls

b=A_1	$n_2 \rightarrow$			
$n_1 \downarrow$	1	1	Φ	Φ
	1	1	Φ	-1
	1	Φ	Φ	-1
	0	Φ	-1	-1
	Φ	Φ	-1	-1
	Φ	-1	-1	-1
	Φ	-1	-1	-1
b=A_2	$n_2 \rightarrow$			
$n_1 \downarrow$	1	0	0	Φ
	1	1	Φ	-1
	1	1	Φ	-1
	1	Φ	-1	-1
	1	Φ	-1	-1
	Φ	-1	-1	-1
	Φ	-1	-1	-1

The simulation and the above analysis has revealed that:

1. Regardless of the numbers of channels reserved, the optimal policy is a control limit policy or threshold policy if the cost functions follow some special prop-

erties. But the numbers of reserved channels do affect the threshold values in the policy.

2. The values of total expected discounted reward $v(s)$ are different if the numbers of reserved channels are changed.

3. If the cost function does not follow the the theory requirement, the optimal policy would not stay in the format of control limit policy.

4. When the reward of a class call is fixed, simply increasing channel capacities does not change the optimal policy for this class calls.

For simple computation, we only did the simulation of 2 classes calls. But by using our algorithm, these could easily be extended to a bigger multimedia network with many classes calls to find the optimal control limit. Also, based on the values of total expected discounted reward $v(s)$ we can design our system to meet some special requirements, like the number of channels that can bring us the best reward when all other parameters are known.

In this chapter we consider the call admission control problem of a single cell with various classes of multimedia traffic under NPS and RCS schemes. We have assumed that arrival processes of calls of each class are Poisson processes, and their service times are exponentially distributed. Admitting each call would bring a reward and there are holding costs per unit time. We find that although the bandwidths, the reserved channels, the arrival/servie rates and rewards/costs for each class are different, the optimal call admission policy for each class of calls is a control limit policy if the holding cost function has some special properties. An example network solution demonstrates the applicability of the theoretical result presented in this paper. The proposed approach and the results derived are applicable in real life for the design and performance analysis of wireless multimedia mobile networks.

Chapter 5

Modeling, Simulation and Analysis of Homogeneous Traffic In a Network

In the past chapters we have considered several models for the traffics on a single cell in a wireless network. In this chapter, we extend our study to the modeling of homogeneous traffic in a wireless network.

5.1 Introduction

To design a wireless cellular network, comparisons need be made between the performance measures of different protocols. Models which provide a multiple cell analysis with an arbitrary mobility pattern must be developed. Consider a wireless network with multiple cells. Assuming that the arrival process of new calls of different classes in each cell is a Poisson process, the call connection time and cell residence time follow exponential distributions, accepting each new call would contribute some units of reward to the system and the system incurs a holding cost per unit time for the calls in the system, we consider the call admission control problem of multiple class calls in a wireless network with multiple cells, and there are routing probabilities between cells. What makes the decision difficult is that, to achieve some sense of optimality, one needs to consider the future status of the network resources and the pattern of the future arrival requests to accept or reject the current incoming calls.

It is well known that call admission control play an important role in the performance of a cellular network. Many different admission control strategies have been discussed in literature [3, 4, 14, 22] to provide priorities to handoff requests without significantly jeopardizing the new connection requests. In the models of this chapter, the handoff calls between cells are always allowed and there could be some channels reserved for the new calls in different cells. We only make the decisions when there is a new call coming into a cell. This is different with the regular channel allocation schemes like Non Priority Scheme, Reserved Channel Scheme (RCS) [23, 24].

5.2 Non Priority Schemes

In this model since handoff calls are always allowed, here Non Priority refer to the new calls in different cells, which means any incoming call to the network is accepted if there are free channels.

5.2.1 Model Formulation

The assumptions and notations are as follows.

1. The network consists of a number (N) of cells.

2. New calls are generated in cell i according to a Poission process with rate λ_i, $i = 1, 2, \ldots, N$. The requested call connection time (RCCT) of a new call at cell i, H_i, is exponentially distributed with means $1/h_i$.

3. The cell residence time, which is defined as the length of time a call stays in the cell and which depends on the velocity and the direction of the mobile terminal, R_i, is exponentially distributed with means $1/r_i$.

4. The probability that a call moves from cell i to a neighboring cell j, given that it moves to a neighboring cell before the call is completed, is $p_{i,j}$, where $\sum_{j=1}^{N} p_{i,j} = 1$. Clearly, $p_{i,i} = 0$ and cell j is a neighboring cell of i if and only if $p_{i,j} > 0$.

5. There are totally C channels in the network. To be fair to each cell, a number of channels k_i is reserved for the new calls in cell i, the handoffs between cells are always allowed.

6. Accepting a new call in cell i would contribute R_i units of reward to the system. Let the number of calls in cells be $\mathbf{n} = (n_1, n_2, \cdots, n_N)$, n_i is the numbers for calls in cell i and the system incurs a holding cost rate $f(\mathbf{n})$ per unit time.

Based on our assumptions, we can build the SMDP model for this wireless mobile network as follows:

1. Let the state variable consists of number of calls in the system, the status of calls leaving, arriving to the system. So $S = \{0, 1, \ldots, C\} \times \{0, 1, \ldots, C\} \times \cdots \times \{0, 1, \ldots, C\} \times \{D, A_1, A_2, \ldots, A_N\}$. In the third set D means a departure from the system, while A_i means an arrival of a new call in cell i. And, $\sum_{i=1}^{N} n_i \leq C$. In general a state could be written as $\langle \mathbf{n}, b \rangle$, where b stands for the last call event, $b \in \{D, A_1, A_2, \ldots, A_N\}$.

2. In states $\langle \mathbf{n}, D \rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D \rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_i \rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_i \rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s,a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b \rangle$ and action a, let $\beta_0 = \sum_{i=1}^{N} (\lambda_i + n_i(r_i + h_i))$, so $\beta(s,a)$ can be written as

$$\beta(s,a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + r_i + h_i, & a = a_A, b = A_i. \end{cases}$$

4. Let $q(z|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle \mathbf{n}, b \rangle$, $b = D, a = a_C$ and $b = A_i, i = (1, 2, \cdots, N), a = a_R$, let $\mathbf{n}_i = \{n_1, n_2, \ldots, \max(n_i - 1, 0), \ldots, n_N\}$ and $\mathbf{n}^i = \{n_1, n_2, \ldots, n_i + 1, \ldots, n_N\}$, $\mathbf{n}_i^j = \{n_1, n_2, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_N\}$, the state transition probabilities are

$$q(z|s,a) = \begin{cases} \lambda_i/\beta_0, & z = \langle \mathbf{n}, A_i \rangle, \\ n_i h_i/\beta_0, & z = \langle \mathbf{n}_i, D \rangle, \\ n_i r_i p_{ij}/\beta_0, & z = \langle \mathbf{n}_i^j, D \rangle. \end{cases}$$

And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_i, i = (1, 2, \cdots, N), a = a_A$, since admitting an incoming call migrates the system state immediately, there is

$$q(z|s, a_A) = q(z|\langle \mathbf{n}^i, D \rangle, a_R).$$

5. Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a)E_s^a\left\{\int_0^t e^{-\alpha\tau}d\tau\right\} \\ &= k(s,a) + c(s,a)E_s^a\left\{[1-e^{-\alpha t}]/\alpha\right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha+\beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \\ R_i, & a = a_A, b = A_i. \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(\mathbf{n}), & a = a_C, a = a_R, \\ -f(\mathbf{n}^i), & a = a_A, b = A_i. \end{cases}$$

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(z|s) = q(z|s,d(s))$, $r_d(s) = r(s,d(s))$ and $\beta_d(s) = \beta(s,d(s))$, we have,

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha+\beta_d(s)}\sum_{z\in S} q_d(z|s)v_\alpha^{d^\infty}(z). \tag{5.1}$$

From equation (5.1), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1-q(s|s,a)]\beta(s,a) \leq c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = C * \sum_{i=1}^{N}(\lambda_i + r_i + h_i)$, we have

$$\tilde{q}(z|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & z = s, \\ \frac{q(z|s,a)\beta(s,a)}{c}, & z \neq s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha + \beta(s,a)}{\alpha + c}.$$

From *Proposition 11.5.1* [13], for each d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{5.2}$$

From equation (5.1) and (5.2), the optimality equation of $v(s)$ for maximum $v_\alpha^\pi(s)$ would have the form of

$$v(s) = \max_{a \in A_s}\left\{\tilde{r}(s,a) + \lambda \sum_{z \in S} \tilde{q}(z|s,a)v(z)\right\}, \tag{5.3}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (5.3), we have for states with $b = D$,

$$\begin{aligned} v(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{i=1}^{N} \lambda_i v(\langle \mathbf{n}, A_i\rangle) \\ + \sum_{i=1}^{N} n_i \mu_i v(\langle \mathbf{n}_i, D\rangle) + (c - \beta_0) v(\langle \mathbf{n}, D\rangle)]. \end{aligned} \tag{5.4}$$

5.2.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_i$,

$$v(\langle \mathbf{n}, A_i \rangle) = \max \left[v(\langle \mathbf{n}, D \rangle), R_i + v(\langle \mathbf{n}^i, D \rangle) \right].$$

Theorem 5.1 *If $v(\langle \boldsymbol{n}, D \rangle)$ is concave nonincreasing on n_i,$i = 1, \ldots, N$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_i(\langle \mathbf{n}, D \rangle) = v(\langle \mathbf{n}^i, D \rangle) - v(\langle \mathbf{n}, D \rangle)$, so $\triangle v_i(\langle \mathbf{n}, D \rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_i, we have the decision rule

$$d(\mathbf{n}) = \begin{cases} a_A, & \triangle v_i(\langle \mathbf{n}, D \rangle) > -R_i, \\ a_R, & \triangle v_i(\langle \mathbf{n}, D \rangle) \leq -R_i. \end{cases}$$

So, if $d(\mathbf{n}) = a_R$, we have $d(\mathbf{n}^i) = a_R$, and so on as i goes through $1, \ldots, N$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 5.2 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_i, i = 1, \ldots, N$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{ij}) \geq f(\boldsymbol{n}^i) + f(\boldsymbol{n}^j),$$

where $\boldsymbol{n}^{ij} = \{n_1, \ldots, n_i{+}1, \ldots, n_j{+}1, \ldots, n_N\}$ and $i, j \in 1, \ldots, N$, then (1): $v(\langle \boldsymbol{n}, D \rangle)$ is concave nonincreasing on $n_i, i = 1, \ldots, N$, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (5.4) to obtain $v^1(\langle \mathbf{n}, D \rangle) = \frac{-f(\mathbf{n})}{\alpha + c}$.

2. Set k=1. We have $v^k(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^k(\langle \mathbf{n}, A_i\rangle)$ is concave nonincreasing on $n_i, i = 1, \ldots, N$. And we have,

$$v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{ij}, b\rangle) \leq v^k(\langle \mathbf{n}^i, b\rangle) + v^k(\langle \mathbf{n}^j, b\rangle),$$

for $b = A_i, i, j \in 1, \ldots, N$.

3. Substituting these v^k back into equation (5.4), we get

$$\begin{aligned} v^{k+1}(\langle \mathbf{n}, D\rangle) &= \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{u=1}^{N} \lambda_u v^k(\langle \mathbf{n}, A_i\rangle) \\ &+ \sum_{u=1}^{N} n_i \mu_u v^k(\langle \mathbf{n}_i, D\rangle) + (c - \beta_0) v^k(\langle \mathbf{n}, D\rangle)]. \end{aligned}$$

so we have for $i = 1, \ldots, N$,

$$\begin{aligned} &v^{k+1}(\langle \mathbf{n}^i, D\rangle) - v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[\\ &\sum_{w=1}^{N} \lambda_w (v^k(\langle \mathbf{n}^i, A_w\rangle) - v^k(\langle \mathbf{n}, A_w\rangle)) \\ &+ \sum_{w=1}^{N} n_w \mu_w (v^k(\langle \mathbf{n}_w^u, D\rangle) - v^k(\langle \mathbf{n}_w, D\rangle)) \\ &+(c - \beta_i)(v^k(\langle \mathbf{n}^i, D\rangle) - v^k(\langle \mathbf{n}, D\rangle)) \\ &+(f(\mathbf{n}) - f(\mathbf{n}^i))], \end{aligned}$$

where $\beta_i = \beta_0 + r_i + h_i$. Since $v^k(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_i, the combination of these functions $v^{k+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_i.

Also, using equation (5.4),

$$\begin{aligned} &v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{ij}, D\rangle) \\ &-v^{k+1}(\langle \mathbf{n}^i, D\rangle) - v^{k+1}(\langle \mathbf{n}^j, D\rangle) \end{aligned}$$

could be divided into several different parts of $v^k(\langle \mathbf{n}, b\rangle)+v^k(\langle \mathbf{n}^{ij}, b\rangle)-v^k(\langle \mathbf{n}^i, b\rangle)-v^k(\langle \mathbf{n}^j, b\rangle), b = D, A_i,\ u = 1, \ldots, N$ which are all ≤ 0. So

$$
\begin{aligned}
& v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{ij}, D\rangle) \\
\leq\ & v^{k+1}(\langle \mathbf{n}^i, D\rangle) + v^{k+1}(\langle \mathbf{n}^j, D\rangle).
\end{aligned}
$$

4. Set k=k+1, go back to step 2.

5. As the iteration continues, $v^k(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^k(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 5.1, the optimal policy is a control limit policy. The verification is over.

5.2.3 Simulation Results

Consider a simple case. Suppose there are only $N = 3$ cells in a wireless cellular network. The routing probabilities between cells are set in Table 5.1. For calculation simplification, let the total channels be $C = 4$, the cost function be $f(\mathbf{n}) = 2n_1^2 + 1.5n_2^2 + n_3^2$, the discount factor $\alpha = 0.1$, and we set the other parameters for analysis as in Table 5.2 to study the performance of the RCS scheme. We get the actions and $v(s)$ values for states in the Table 5.3 and Table 5.4.

Table 5.1: Routing Probablities

Cell	1	2	3
1	0	0.4	0.6
2	0.5	0	0.5
3	0.8	0.2	0

In Table 5.2, λ is the arrival rate of new calls, h is the call connection rate, r is the cell residence rate and R is the reward for new calls.

Table 5.2: Parameters

Cell	λ	h	r	R
1	4	6	4	2
2	2	4	5	1.6
3	1	3	2	1.5

Table 5.3: Actions for calls when $n_3 = 0$

$b = A_1$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	1	0
	1	1	1	0	-1
	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_2$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	1	0
	1	1	1	0	-1
	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_3$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	1	0
	1	1	1	0	-1
	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1

Table 5.3 shows the actions for calls when $n_3 = 0$, which means there is currently no call in cell 3. In Table 5.3, '1' means to admit a new call, '0' are those states in which the system does not accept any new calls. '-1' denotes those states that do not exist in the state space. As seen from Table 5.3, all the arrival of new calls are allowed. It is seen that when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the new calls in the cells, so the optimal policy is a control limit policy, which is consistent with Theorem 5.2.

Table 5.4: $v(s)$ for states

b=D	$n_2 \rightarrow$				
$n_1 \downarrow$	73.6962	73.0631	72.2104	71.1140	69.6937
	72.8953	72.1625	71.1878	69.8997	0
	71.8124	70.9596	69.8043	0	0
	70.4296	69.4075	0	0	0
	68.7096	0	0	0	0
b=A_1	$n_2 \rightarrow$				
$n_1 \downarrow$	74.8953	74.1625	73.1878	71.8997	69.6937
	73.8124	72.9596	71.8043	69.8997	0
	72.4296	71.4075	69.8043	0	0
	70.7096	69.4075	0	0	0
	68.7096	0	0	0	0
b=A_2	$n_2 \rightarrow$				
$n_1 \downarrow$	74.6631	73.8104	72.7140	71.2937	69.6937
	73.7625	72.7878	71.4997	69.8997	0
	72.5596	71.4043	69.8043	0	0
	71.0075	69.4075	0	0	0
	68.7096	0	0	0	0
b=A_3	$n_2 \rightarrow$				
$n_1 \downarrow$	74.3359	73.6008	72.6185	71.3081	69.6937
	73.3888	72.5293	71.3509	69.8997	0
	72.1382	71.0924	69.8043	0	0
	70.5325	69.4075	0	0	0
	68.7096	0	0	0	0

In Table 5.4, '0' denotes the states that do not exist in the state space. As for this special case $n_3 = 0$, we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion of Theorem 5.2.

5.3 Reserved Channel Schemes

In this model the handoff calls are always allowed, but we can reserve some channels for the new calls in different cells and still under the name of 'Reserved Channel Scheme'.

5.3.1 Model Formulation

The assumptions and notations are as follows.

1. The network consists of a number (N) of cells.

2. New calls are generated in cell i according to a Poission process with rate λ_i, $i = 1, 2, \ldots, N$. The requested call connection time (RCCT) of a new call at cell i, H_i, is exponentially distributed with means $1/h_i$.

3. The cell residence time, which is defined as the length of time a call stays in the cell and which depends on the velocity and the direction of the mobile terminal, R_i, is exponentially distributed with means $1/r_i$.

4. The probability that a call moves from cell i to a neighboring cell j, given that it moves to a neighboring cell before the call is completed, is $p_{i,j}$, where $\sum_{j=1}^{N} p_{i,j} = 1$. Clearly, $p_{i,i} = 0$ and cell j is a neighboring cell of i if and only if $p_{i,j} > 0$.

5. There are totally C channels in the network. To be fair to each cell, a number of channels k_i is reserved for the new calls in cell i, the handoffs between cells are always allowed.

6. Accepting a new call in cell i would contribute R_i units of reward to the system. Let the number of calls in cells be $\mathbf{n} = (n_1, n_2, \cdots, n_N)$, n_i is the numbers for calls in cell i and the system incurs a holding cost rate $f(\mathbf{n})$ per unit time.

Based on our assumptions, we can build the SMDP model for this wireless mobile network as follows:

1. Let the state variable consists of number of calls in the system, the status of calls leaving, arriving to the system. So $S = \{0, 1, \ldots, C\} \times \{0, 1, \ldots, C\} \times \cdots \times \{0, 1, \ldots, C\} \times \{D, A_1, A_2, \ldots, A_N\}$. In the third set D means a departure from

the system, while A_i means an arrival of a new call in cell i. And, $\sum_{i=1}^{N} n_i \leq C$. In general a state could be written as $\langle \mathbf{n}, b \rangle$, where b stands for the last call event, $b \in \{D, A_1, A_2, \ldots, A_N\}$.

2. In states $\langle \mathbf{n}, D \rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D \rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_i \rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_i \rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s, a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b \rangle$ and action a, let $\beta_0 = \sum_{i=1}^{N} (\lambda_i + n_i(r_i + h_i))$, so $\beta(s, a)$ can be written as

$$\beta(s, a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + r_i + h_i, & a = a_A, b = A_i. \end{cases}$$

4. Let $q(z|s, a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle \mathbf{n}, b \rangle$, $b = D, a = a_C$ and $b = A_i, i = (1, 2, \cdots, N), a = a_R$, let $\mathbf{n}_i = \{n_1, n_2, \ldots, \max(n_i - 1, 0), \ldots, n_N\}$ and $\mathbf{n}^i = \{n_1, n_2, \ldots, n_i + 1, \ldots, n_N\}$, $\mathbf{n}_i^j = \{n_1, n_2, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_N\}$,

the state transition probabilities are

$$q(z|s,a) = \begin{cases} \lambda_i/\beta_0, & z = \langle \mathbf{n}, A_i \rangle, \\ n_i h_i/\beta_0, & z = \langle \mathbf{n}_i, D \rangle, \\ n_i r_i p_{ij}/\beta_0, & z = \langle \mathbf{n}_i^j, D \rangle. \end{cases}$$

And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_i, i = (1, 2, \cdots, N), a = a_A$, since admitting an incoming call migrates the system state immediately, there is

$$q(z|s, a_A) = q(z|\langle \mathbf{n}^i, D \rangle, a_R).$$

5. Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a) E_s^a \left\{ \int_0^t e^{-\alpha\tau} d\tau \right\} \\ &= k(s,a) + c(s,a) E_s^a \left\{ [1 - e^{-\alpha t}]/\alpha \right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha + \beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \\ R_i, & a = a_A, b = A_i. \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(\mathbf{n}), & a = a_C, a = a_R, \\ -f(\mathbf{n}^i), & a = a_A, b = A_i. \end{cases}$$

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(z|s) = q(z|s,d(s))$, $r_d(s) = r(s,d(s))$ and $\beta_d(s) = \beta(s,d(s))$, we have,

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{z \in S} q_d(z|s) v_\alpha^{d^\infty}(z). \tag{5.5}$$

From equation (5.5), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1 - q(s|s,a)]\beta(s,a) \leq c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = C * \sum_{i=1}^{N}(\lambda_i + r_i + h_i)$, we have

$$\tilde{q}(z|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & z = s, \\ \frac{q(z|s,a)\beta(s,a)}{c}, & z \neq s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha + \beta(s,a)}{\alpha + c}.$$

From *Proposition 11.5.1* [13], for each d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{5.6}$$

From equation (5.5) and (5.6), the optimality equation of $v(s)$ for maximum $v_\alpha^\pi(s)$ would have the form of

$$v(s) = \max_{a \in A_s} \left\{ \tilde{r}(s,a) + \lambda \sum_{z \in S} \tilde{q}(z|s,a) v(z) \right\}, \tag{5.7}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (5.7), we have for states with $b = D$,

$$\begin{aligned} v(\langle \mathbf{n}, D \rangle) = & \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{i=1}^{N} \lambda_i v(\langle \mathbf{n}, A_i \rangle) \\ & + \sum_{i=1}^{N} n_i \mu_i v(\langle \mathbf{n}_i, D \rangle) + (c - \beta_0) v(\langle \mathbf{n}, D \rangle)]. \end{aligned} \tag{5.8}$$

5.3.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_i$,

$$v(\langle \mathbf{n}, A_i \rangle) = \max \left[v(\langle \mathbf{n}, D \rangle), R_i + v(\langle \mathbf{n}^i, D \rangle) \right].$$

Theorem 5.3 *If $v(\langle \boldsymbol{n}, D \rangle)$ is concave nonincreasing on n_i,$i = 1, \ldots, N$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_i(\langle \mathbf{n}, D \rangle) = v(\langle \mathbf{n}^i, D \rangle) - v(\langle \mathbf{n}, D \rangle)$, so $\triangle v_i(\langle \mathbf{n}, D \rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_i, we have the decision rule

$$d(\mathbf{n}) = \begin{cases} a_A, & \triangle v_i(\langle \mathbf{n}, D \rangle) > -R_i, \\ a_R, & \triangle v_i(\langle \mathbf{n}, D \rangle) \leq -R_i. \end{cases}$$

So, if $d(\mathbf{n}) = a_R$, we have $d(\mathbf{n}^i) = a_R$, and so on as i goes through $1, \ldots, N$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 5.4 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_i, i = 1, \ldots, N$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{ij}) \geq f(\boldsymbol{n}^i) + f(\boldsymbol{n}^j),$$

where $\boldsymbol{n}^{ij} = \{n_1, \ldots, n_i{+}1, \ldots, n_j{+}1, \ldots, n_N\}$ and $i, j \in 1, \ldots, N$, then (1): $v(\langle \boldsymbol{n}, D\rangle)$ is concave nonincreasing on $n_i, i = 1, \ldots, N$, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (5.8) to obtain $v^1(\langle \mathbf{n}, D\rangle) = \frac{-f(\mathbf{n})}{\alpha+c}$.

2. Set k=1. We have $v^k(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^k(\langle \mathbf{n}, A_i\rangle)$ is concave nonincreasing on $n_i, i = 1, \ldots, N$. And we have,

$$v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{ij}, b\rangle) \leq v^k(\langle \mathbf{n}^i, b\rangle) + v^k(\langle \mathbf{n}^j, b\rangle),$$

   for $b = A_i, i, j \in 1, \ldots, N$.

3. Substituting these v^k back into equation (5.8), we get

$$v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{u=1}^{N} \lambda_u v^k(\langle \mathbf{n}, A_i\rangle)$$
$$+ \sum_{u=1}^{N} n_i \mu_u v^k(\langle \mathbf{n}_i, D\rangle) + (c - \beta_0) v^k(\langle \mathbf{n}, D\rangle)].$$

   so we have for $i = 1, \ldots, N$,

$$v^{k+1}(\langle \mathbf{n}^i, D\rangle) - v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[$$
$$\sum_{w=1}^{N} \lambda_w (v^k(\langle \mathbf{n}^i, A_w\rangle) - v^k(\langle \mathbf{n}, A_w\rangle))$$

$$
\begin{aligned}
&+\sum_{w=1}^{N} n_w\mu_w(v^k(\langle \mathbf{n}_w^u, D\rangle) - v^k(\langle \mathbf{n}_w, D\rangle)) \\
&+(c-\beta_i)(v^k(\langle \mathbf{n}^i, D\rangle) - v^k(\langle \mathbf{n}, D\rangle)) \\
&+(f(\mathbf{n}) - f(\mathbf{n}^i))],
\end{aligned}
$$

where $\beta_i = \beta_0 + r_i + h_i$. Since $v^k(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_i, the combination of these functions $v^{k+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_i.

Also, using equation (5.8),

$$
\begin{aligned}
&v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{ij}, D\rangle) \\
&-v^{k+1}(\langle \mathbf{n}^i, D\rangle) - v^{k+1}(\langle \mathbf{n}^j, D\rangle)
\end{aligned}
$$

could be divided into several different parts of $v^k(\langle \mathbf{n}, b\rangle)+v^k(\langle \mathbf{n}^{ij}, b\rangle)-v^k(\langle \mathbf{n}^i, b\rangle)-v^k(\langle \mathbf{n}^j, b\rangle), b = D, A_i,\ u = 1, \ldots, N$ which are all ≤ 0. So

$$
\begin{aligned}
&v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{ij}, D\rangle) \\
&\leq v^{k+1}(\langle \mathbf{n}^i, D\rangle) + v^{k+1}(\langle \mathbf{n}^j, D\rangle).
\end{aligned}
$$

4. Set k=k+1, go back to step 2.

5. As the iteration continues, $v^k(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^k(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 5.3, the optimal policy is a control limit policy. The verification is over.

5.3.3 Simulation Results

Consider a simple case. Suppose there are only $N = 3$ cells in this wireless network. The routing probabilities between cells are set in Table 5.5. For simplification, let the total channels be $C = 4$, the cost function be $f(\mathbf{n}) = 2n_1^2 + 1.5n_2^2 + n_3^2$, the discount factor $\alpha = 0.1$, and we set the other parameters for analysis as in Table 5.6 to study the performance of the RCS scheme. We get the actions and $v(s)$ values for states in the Table 5.7 when $n_3 = 0$.

Table 5.5: Routing Probabilities

Cell	1	2	3
1	0	0.4	0.6
2	0.5	0	0.5
3	0.8	0.2	0

Table 5.6: Parameters

Cell	λ	h	r	R	K
1	4	6	4	2	1
2	2	4	5	1.6	0
3	1	3	2	1.5	0

In Table 5.6, λ is the arrival rate of new calls, h is the call connection rate, r is the cell residence rate and R is the reward for new calls, and K is the number of reserved channels.

Table 5.7 shows the actions for calls when $n_3 = 0$, which means there is currently no call in cell 3. In Table 5.7, '1' is to accept the new call, '0' are those states in which the system does not accept any new calls. '-1' denotes those states that do not exist in the state space. As seen from Table 5.7, because of the reserved channel in cell 1 and its bigger reward, only the arrival of new calls in cell 1 is always allowed. Because all the handoff calls are allowed in the system, although there is one channel reserved for the new calls in cell 1, the system can still have $n_2 = 4$, this is due to the

Table 5.7: Actions for calls when $n_3 = 0$

$b = A_1$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	1	0
	1	1	1	0	-1
	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_2$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	0	0
	1	1	0	0	-1
	1	0	0	-1	-1
	0	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_3$	$n_2 \rightarrow$				
$n_1 \downarrow$	1	1	1	0	0
	1	0	0	0	-1
	0	0	0	-1	-1
	0	0	-1	-1	-1
	0	-1	-1	-1	-1

handoff calls routing from other cells to cell 2. It is seen that when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the new calls in the cells, so the optimal policy is a control limit policy, which is consistent with Theorem 5.4.

Table 5.8: $v(s)$ for states

b=D	$n_2 \rightarrow$				
$n_1 \downarrow$	71.8344	71.1799	70.2980	69.1600	67.7396
	71.0064	70.2412	69.2164	67.9352	0
	69.8816	68.9701	67.8298	0	0
	68.4211	67.4233	0	0	0
	66.7160	0	0	0	0
b=A_1	$n_2 \rightarrow$				
$n_1 \downarrow$	73.0064	72.2412	71.2164	69.9352	67.7396
	71.8816	70.9701	69.8298	67.9352	0
	70.4211	69.4233	67.8298	0	0
	68.7160	67.4233	0	0	0
	66.7160	0	0	0	0
b=A_2	$n_2 \rightarrow$				
$n_1 \downarrow$	72.7799	71.8980	70.7600	69.1600	67.7396
	71.8412	70.8164	69.2164	67.9352	0
	70.5701	68.9701	67.8298	0	0
	68.4211	67.4233	0	0	0
	66.7160	0	0	0	0
b=A_3	$n_2 \rightarrow$				
$n_1 \downarrow$	72.4393	71.6716	70.6442	69.1600	67.7396
	71.4494	70.5351	69.2164	67.9352	0
	70.1232	68.9701	67.8298	0	0
	68.4211	67.4233	0	0	0
	66.7160	0	0	0	0

In Table 5.8, '0' denotes the states that do not exist in the state space. As for this special case $n_3 = 0$, we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion of Theorem 5.4.

5.4 Numerical Analysis

As seen from the previous sections, to compare the NPS and RCS schemes, we set the same parameters and routing probabilities among the cells, the only difference is on the number of channels reserved.

Compare the actions in Table 5.3 and Table 5.7, based on the reward, the actions with NPS scheme is to accept any incoming arrival calls. But with the RCS scheme, due to reserved channel for the new call in cell 1, the arrival new calls in cell 2 and 3 can not be always accepted. Also the difference is shown in the Table 5.4 and Table 5.8 for the $v(s)$ values.

We have shown the actions of $n_3 = 0$, using the same parameter setting, We show the actions and $v(s)$ values for states in the when $n_1 = 0$ under NPS and RCS schemes.

Table 5.9: Actions for calls when $n_1 = 0$, NPS

$b = A_1$	$n_3 \rightarrow$				
	1	1	1	1	0
	1	1	1	0	-1
$n_2 \downarrow$	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_2$	$n_3 \rightarrow$				
	1	1	1	1	0
	1	1	1	0	-1
$n_2 \downarrow$	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_3$	$n_3 \rightarrow$				
	1	1	1	1	0
	1	1	0	0	-1
$n_2 \downarrow$	1	0	0	-1	-1
	0	0	-1	-1	-1
	0	-1	-1	-1	-1

It is seen from 5.3 that the new arrivals in cell 1 and 2 are always allowed, but due to the smaller reward on accepting the new calls in cell 3, the A_3 arrival is rejected sometimes.

Compare the actions in Table 5.9 and Table 5.10, because of the 1 channel reserved for the A_1, A_1 arrival is always allowed, but not the same to A_2 and A_3 arrivals.

The simulation and the above analysis has revealed that:

1. Regardless of the numbers of channels reserved, the optimal policy is a control limit policy or threshold policy if the cost functions follow some special prop-

Table 5.10: Actions for calls when $n_1 = 0$, RCS

$b = A_1$	$n_3 \rightarrow$				
$n_2 \downarrow$	1	1	1	1	0
	1	1	1	0	-1
	1	1	0	-1	-1
	1	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_2$	$n_3 \rightarrow$				
$n_2 \downarrow$	1	1	1	0	0
	1	1	0	0	-1
	0	1	0	-1	-1
	0	0	-1	-1	-1
	0	-1	-1	-1	-1
$b = A_3$	$n_3 \rightarrow$				
$n_2 \downarrow$	1	1	1	0	0
	1	1	0	0	-1
	0	0	0	-1	-1
	0	0	-1	-1	-1
	0	-1	-1	-1	-1

erties. But the numbers of reserved channels do affect the threshold values in the policy.

2. The values of total expected discounted reward $v(s)$ are different if the numbers of reserved channels are changed.

3. When the reward of a class call is fixed, simply increasing channel capacities does not change the optimal policy for this class calls.

For simple computation, we only did the simulation of a wireless network with 3 cells. But by using our algorithm, these could easily be extended to a bigger network to find the optimal control limit. Also, based on the values of total expected discounted reward $v(s)$ we can design our system to meet some special requirements, like the number of channels that can bring us the best reward when all other parameters are known.

In this chapter, we consider the homogeneous traffic in a wireless network with multiple cells and assumed the routing probabilities between cells. Also assume that the arrival process of new calls in each cell is a Poisson process, the call connection time and cell residence time follow exponential distributions, accepting each new call would contribute some units of reward to the system and the system incurs a holding cost per unit time for the calls in the system. We find that if the holding cost function is convex nondecreasing, the optimal policy call admission policy for calls is a control limit policy. The numerical analysis has proven our theoretical conclusion and described the relationship between parameters. The result of this paper could easily be used in designing wireless networks for optimal performance.

Chapter 6

Modeling, Simulation and Analysis of Multimedia Traffic In a Network

In the past chapters we have considered several models for the traffics on a single cell in a wireless network and homogeneous traffic in a wireless network. In this chapter, we consider multimedia traffic in a network.

6.1 Introduction

Modeling and analysis of telecommunication networks have been the subject of extensive studies for a long time, [7–9, 11, 17]. With the advent of the third generation of wireless multimedia services, brought about the need to adapt the existing mobile cellular networks to make them carry the various classes of multimedia traffic (voice, video, images, web documents, data and a combination thereof). The wireless multimedia network has to be able to support multiple classes of traffic with different Quality of Service (QoS) requirements, i.e., different number of channels needed, holding time of the connection and cell residence time.

To design a wireless cellular network, comparisons need to be made between the performance measures of different protocols. Models which provide a multiple cell analysis with an arbitrary mobility pattern must be developed. In wireless networks, the calls are normally divided into two groups: new calls and handoff calls. When a user moves from one cell to another, the base station in the new cell must take responsibility for all the previously established connections. Since premature termination of established connections is usually more objectionable than rejection of a new connection request, it is widely believed that a wireless network must give higher priority to the handoff connection requests as compared to new connection requests.

In the models of this chapter, seem as in the last chapter, the handoff calls between cells are always allowed and there could be some channels reserved for the new calls in different cells. We only make the decisions when there is a new call coming into a cell. This is different with the regular channel allocation schemes like Non Priority Scheme, Reserved Channel Scheme (RCS) [23, 24].

6.2 Non Priority Schemes

For every class of calls, assume that the new traffic origination rates are uniformly distributed over the mobile service area. Any incoming call to the cell can get a service if there are enough free channels in the system.

6.2.1 Model Formulation

Consider there are U classes of multimedia calls (telephone, video, etc., but for convenience we shall call all of them *calls*) in the network. The other assumptions and notations for this wireless network are as follows.

1. The required bandwidth of class u calls $u = 1, \ldots, U$ is b_u.
2. The network consists of a number (N) of cells.
3. New calls of class u are generated in cell i according to a Poission process with

rate λ_{ui}, $i = 1, 2, \ldots, N$. The requested call connection time (RCCT) of a new call at cell i, H_{ui}, is exponentially distributed with means $1/h_{ui}$.

4. The cell residence time, which is defined as the length of time a class u call stays in the cell i and which depends on the velocity and the direction of the mobile terminal, R_{ui}, is exponentially distributed with means $1/r_{ui}$.

5. The probability that a call moves from cell i to a neighboring cell j, given that it moves to a neighboring cell before the call is completed, is $p_{i,j}$, where $\sum_{j=1}^{N} p_{i,j} = 1$. Clearly, $p_{i,i} = 0$ and cell j is a neighboring cell of i if and only if $p_{i,j} > 0$.

6. There are a total of C channels in the network. A number of channels K_{ui} can be reserved for the new calls of class u in cell i if necessary, and the handoff calls among cells are always allowed.

7. Accepting a new call of class u in cell i would contribute R_{ui} units of reward to the system. Let the number of calls in cells be
$\mathbf{n} = (n_{11}, \ldots, n_{U1}, n_{12}, \ldots, n_{U2}, \cdots, n_{1N}, \ldots, n_{UN})$,
n_{ui} is the numbers for calls of class u in cell i and the system incurs a holding cost rate $f(\mathbf{n})$ per unit time.

Based on these assumptions, we can build the SMDP model for this wireless cellular network as follows:

1. Let the state variable consists of number of calls in the system, the status of calls leaving, arriving to the system. So
$S = \{0, 1, \ldots, \lfloor \frac{C}{b_1} \rfloor\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C}{b_U} \rfloor\} \times \cdots \times \{0, 1, \ldots, \lfloor \frac{C}{b_1} \rfloor\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C}{b_U} \rfloor\} \times \{D, A_{11}, \ldots, A_{U1}, A_{12}, \ldots, A_{U2}, \cdots, A_{1N}, \ldots, A_{UN}\}$, where $\lfloor x \rfloor$ denotes the nearest integer to x. In the third set D means a departure from the system, while A_{ui} means an arrival of a new call of class u in cell i. And,

$\sum_{u=1}^{U}\sum_{i=1}^{N} n_{ui}b_u \leq C$. In general a state could be written as $\langle \mathbf{n}, b\rangle$, where b stands for the last call event, $b \in \{D, A_{11}, \ldots, A_{U1}, A_{12}, \ldots, A_{U2}, \cdots, A_{1N}, \ldots, A_{UN}\}$.

2. In states $\langle \mathbf{n}, D\rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D\rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_{ui}\rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_{ui}\rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s,a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b\rangle$ and action a, let $\beta_0 = \sum_{u=1}^{U}\sum_{i=1}^{N}(\lambda_{ui} + n_{ui}(r_{ui} + h_{ui}))$, so $\beta(s,a)$ can be written as

$$\beta(s,a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + r_{ui} + h_{ui}, & a = a_A, b = A_{ui}. \end{cases}$$

where $u = (1, 2, \cdots, U), i = (1, 2, \cdots, N)$.

4. Let $q(z|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle \mathbf{n}, b\rangle$, $b = D, a = a_C$ and $b = A_{ui}, a = a_R$, let

$\mathbf{n}_{ui} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots \max(n_{ui} - 1, 0) \ldots, n_{Ui}, \cdots, n_{1N}, \ldots, n_{UN}\}$

and

$\mathbf{n}^{ui} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots, n_{ui} + 1, \ldots, n_{Ui}, \cdots, n_{1N}, \ldots, n_{UN}\}$,

$\mathbf{n}_{ui}^{uj} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots, n_{ui} - 1, \ldots, n_{Ui}, \cdots, n_{1j}, \cdots, n_{uj} + 1, \ldots, n_{Uj},$

$\cdots, n_{1N}, \ldots, n_{UN}\}$,

the state transition probabilities are

$$q(z|s,a) = \begin{cases} \lambda_{ui}/\beta_0, & z = \langle \mathbf{n}, A_{ui} \rangle, \\ n_{ui}h_{ui}/\beta_0, & z = \langle \mathbf{n}_{ui}, D \rangle, \\ n_{ui}r_{ui}p_{ij}/\beta_0, & z = \langle \mathbf{n}_{ui}^{uj}, D \rangle. \end{cases}$$

And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_{ui}, u = (1, 2, \cdots, U), i = (1, 2, \cdots, N), a = a_A$, since admitting an incoming call migrates the system state immediately, there is

$$q(z|s, a_A) = q(z|\langle \mathbf{n}^{ui}, D \rangle, a_R).$$

5. Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

$$\begin{aligned} r(s,a) &= k(s,a) + c(s,a)E_s^a \left\{ \int_0^t e^{-\alpha\tau} d\tau \right\} \\ &= k(s,a) + c(s,a)E_s^a \left\{ [1 - e^{-\alpha t}]/\alpha \right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha + \beta(s,a)}, \end{aligned}$$

where

$$k(s,a) = \begin{cases} 0, & a = a_C, a = a_R, \\ R_{ui}, & a = a_A, b = A_{ui}. \end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action

a. We have the holding cost rate as

$$c(s,a) = \begin{cases} -f(\mathbf{n}), & a = a_C, a = a_R, \\ -f(\mathbf{n}^{ui}), & a = a_A, b = A_{ui}. \end{cases}$$

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(z|s) = q(z|s,d(s))$, $r_d(s) = r(s,d(s))$ and $\beta_d(s) = \beta(s,d(s))$, we have,

$$v_\alpha^{d^\infty}(s) = r_d(s) + \frac{\beta_d(s)}{\alpha + \beta_d(s)} \sum_{z \in S} q_d(z|s) v_\alpha^{d^\infty}(z). \tag{6.1}$$

From equation (6.1), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1 - q(s|s,a)]\beta(s,a) \le c, \forall s \in S, a \in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S} = S$, $\tilde{A}_s = A_s$, and $c = C * \sum_{i=1}^{N}(\lambda_i + r_i + h_i)$, we have

$$\tilde{q}(z|s,a) = \begin{cases} 1 - \frac{[1-q(s|s,a)]\beta(s,a)}{c}, & z = s, \\ \frac{q(z|s,a)\beta(s,a)}{c}, & z \neq s. \end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha + \beta(s,a)}{\alpha + c}.$$

From *Proposition 11.5.1* [13], for each d^∞ policy and $s \in S$, we have

$$\tilde{v}_\alpha^{d^\infty}(s) = v_\alpha^{d^\infty}(s). \tag{6.2}$$

From equation (6.1) and (6.2), the optimality equation of $v(s)$ for maximum $v_\alpha^\pi(s)$ would have the form of

$$v(s) = \max_{a \in A_s} \left\{ \tilde{r}(s,a) + \lambda \sum_{z \in S} \tilde{q}(z|s,a)v(z) \right\}, \tag{6.3}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (6.3), we have for states with $b = D$,

$$\begin{aligned} v(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[&-f(\mathbf{n}) + \sum_{u=1}^{U}\sum_{i=1}^{N}\lambda_{ui} v(\langle \mathbf{n}, A_{ui}\rangle) \\ &+ \sum_{u=1}^{U}\sum_{i=1}^{N} n_{ui}h_{ui}v(\langle \mathbf{n}_{ui}, D\rangle) + \sum_{u=1}^{U}\sum_{j=1}^{N}\sum_{i=1}^{N} p_{ij}n_{ui}r_{ui}v(\langle \mathbf{n}_{ui}^{uj}, D\rangle) \\ &+ (c - \beta_0)v(\langle \mathbf{n}, D\rangle)]. \end{aligned} \tag{6.4}$$

6.2.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_{ui}$,

$$v(\langle \mathbf{n}, A_{ui}\rangle) = \max\left[v(\langle \mathbf{n}, D\rangle), R_{ui} + v(\langle \mathbf{n}^{ui}, D\rangle)\right].$$

Theorem 6.1 *If $v(\langle \boldsymbol{n}, D\rangle)$ is concave nonincreasing on n_{ui}, $u = 1,\ldots,U, i = 1,\ldots,N$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_{ui}(\langle \mathbf{n}, D\rangle) = v(\langle \mathbf{n}^{ui}, D\rangle) - v(\langle \mathbf{n}, D\rangle)$, so $\triangle v_{ui}(\langle \mathbf{n}, D\rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_{ui}, we have the decision rule

$$d(\mathbf{n}) = \begin{cases} a_A, & \triangle v_{ui}(\langle \mathbf{n}, D\rangle) > -R_{ui}, \\ a_R, & \triangle v_{ui}(\langle \mathbf{n}, D\rangle) \leq -R_{ui}. \end{cases}$$

So, if $d(\mathbf{n}) = a_R$, we have $d(\mathbf{n}^{ui}) = a_R$, and so on as u goes through $1, \ldots, U$ and i goes through $1, \ldots, N$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 6.2 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_{ui}, u = 1, \ldots, U, i = 1, \ldots, N$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{uiuj}) \geq f(\boldsymbol{n}^{ui}) + f(\boldsymbol{n}^{uj}),$$

where $\boldsymbol{n}^{uiuj} = \{n_{11}, \ldots, n_{U1}, n_{12}, \ldots, n_{U2}, \cdots, n_{1i}, \ldots n_{ui} + 1 \ldots, n_{Ui}, \cdots, n_{1j}, \ldots n_{uj} + 1 \ldots, n_{Uj}, \cdots, n_{1N}, \ldots, n_{UN}\}$ and $i, j \in 1, \ldots, N$, then (1): $v(\langle \boldsymbol{n}, D\rangle)$ is concave nonincreasing on $n_{ui}, i = 1, \ldots, N$, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (6.4) to obtain $v^1(\langle \mathbf{n}, D\rangle) = \frac{-f(\mathbf{n})}{\alpha + c}$.

2. Set k=1. We have $v^k(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^k(\langle \mathbf{n}, A_{ui}\rangle)$ is concave nonincreasing on $n_{ui}, u = 1, \ldots, U, i = 1, \ldots, N$. And we have,

$$v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{uiuj}, b\rangle) \leq v^k(\langle \mathbf{n}^{ui}, b\rangle) + v^k(\langle \mathbf{n}^{uj}, b\rangle),$$

   for $b = A_{ui}, u = 1, \ldots, U, \& i, j \in 1, \ldots, N$.

3. Substituting these v^k back into equation (6.4), we get

$$
\begin{aligned}
v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha+c}[-f(\mathbf{n}) + \sum_{u=1}^{U}\sum_{i=1}^{N}\lambda_{ui}v^k(\langle \mathbf{n}, A_{ui}\rangle) \\
+\sum_{u=1}^{U}\sum_{i=1}^{N} n_{ui}h_{ui}v^k(\langle \mathbf{n}_{ui}, D\rangle) + \sum_{u=1}^{U}\sum_{j=1}^{N}\sum_{i=1}^{N} p_{ij}n_{ui}r_{ui}v^k(\langle \mathbf{n}_{ui}^{uj}, D\rangle) \\
+(c-\beta_0)v^k(\langle \mathbf{n}, D\rangle)].
\end{aligned}
$$

so we have for $u = 1, \ldots, U, i = 1, \ldots, N$,

$$
\begin{aligned}
v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) - v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha+c}[\\
\sum_{v=1}^{U}\sum_{w=1}^{N}\lambda_{vw}(v^k(\langle \mathbf{n}^{ui}, A_{vw}\rangle) - v^k(\langle \mathbf{n}, A_{vw}\rangle)) \\
+\sum_{v=1}^{U}\sum_{w=1}^{N} n_{vw}h_{vw}(v^k(\langle \mathbf{n}_{vw}^{ui}, D\rangle) - v^k(\langle \mathbf{n}_{vw}, D\rangle)) \\
+\sum_{v=1}^{U}\sum_{j=1}^{N}\sum_{w=1}^{N} p_{wj}n_{vw}r_{vw}(v^k(\langle \mathbf{n}_{vw}^{uivj}, D\rangle) - v^k(\langle \mathbf{n}_{vw}^{vj}, D\rangle)) \\
+(c-\beta_{ui})(v^k(\langle \mathbf{n}^{ui}, D\rangle) - v^k(\langle \mathbf{n}, D\rangle)) \\
+(f(\mathbf{n}) - f(\mathbf{n}^{ui}))],
\end{aligned}
$$

where $\beta_i = \beta_0 + r_{ui} + h_{ui}$. Since $v^k(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_i, the combination of these functions $v^{k+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_i.

Also, using equation (6.4),

$$
\begin{aligned}
v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uiuj}, D\rangle) \\
-v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) - v^{k+1}(\langle \mathbf{n}^{uj}, D\rangle)
\end{aligned}
$$

could be divided into several different parts of $v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{uiuj}, b\rangle) -$

$v^k(\langle \mathbf{n}^{ui}, b\rangle) - v^k(\langle \mathbf{n}^{uj}, b\rangle)$, where $b = D, A_{ui}$, which are all ≤ 0. So

$$v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uiuj}, D\rangle)$$
$$\leq v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uj}, D\rangle).$$

4. Set k=k+1, go back to step 2.

5. As the iteration continues, $v^k(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^k(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 6.1, the optimal policy is a control limit policy. The verification is over.

6.2.3 Simulation Results

Without losing generosity, suppose there are $N = 3$ cells and 2 classes of traffic in this wireless network. The routing probabilities between cells are set in Table 6.1. For simplification, let the total channels be $C = 6$, the cost function be $f(\mathbf{n}) = 2n_{11}^2 + 1.5n_{12}^2 + n_{13}^2 + n_{21}^2 + n_{22}^2 + 0.5n_{23}^2$, the discount factor $\alpha = 0.1$, the bandwidths for two classes are $b_1 = 1, b_2 = 2$, and we set the other parameters for analysis as in Table 6.2 to study the performance of the RCS scheme. We get the actions and $v(s)$ values for states in the Table 6.3 and Table 6.4.

Table 6.1: Routing Probablities

Cell	1	2	3
1	0	0.4	0.6
2	0.5	0	0.5
3	0.8	0.2	0

In Table 6.2, λ is the arrival rate of new calls, h is the call connection rate, r is the cell residence rate, R is the reward for new calls.

Table 6.2: Parameters Setting

Class	Cell	λ	h	r	R	K
1	1	4	6	4	5	0
1	2	2	4	5	4	0
1	3	1	3	2	6	1
2	1	4	6	4	8	0
2	2	2	4	5	7	0
2	3	1	3	2	6	1

Table 6.3: Actions for calls arrival in Cell 3

b = A_{13}	n_{12}						
	1	1	1	1	1	1	0
	1	1	1	1	1	0	-1
	1	1	1	1	0	-1	-1
n_{11}	1	1	1	0	-1	-1	-1
	1	1	0	-1	-1	-1	-1
	1	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1
b = A_{23}	n_{12}						
	1	1	1	1	1	0	0
	1	1	1	1	0	0	-1
	1	1	1	0	0	-1	-1
n_{11}	1	1	0	0	-1	-1	-1
	1	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1

Table 6.3 shows the actions for call arrivals when there are no class 2 calls in the system and $n_{13} = 0$, which means there is currently no call of class 1 in cell 3. In Table 6.3, '1' means to admit a new call , and '0' are those states in which the system does not accept any new calls. '-1' denotes those states that do not exist in the state space. As seen from Table 6.3, all the call arrivals can get a service if there are enough free channels, but due to the difference on bandwidth, the actions status for b = A_{13} and b = A_{23} are not totally the same. It is seen that when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the new calls in the cells, so the optimal policy is a control limit policy, which is consistent with Theorem 6.2.

Table 6.4: $v(s)$ for states

b=D	$n_2 \rightarrow$						
$n_1 \downarrow$	521.2774	519.7459	517.9141	515.6880	513.0994	509.8119	506.2738
	519.4093	517.6879	515.5675	513.0891	509.9056	506.4910	0
	517.1581	515.1429	512.7752	509.6954	506.4050	0	0
	514.4142	512.1575	509.1813	506.0160	0	0	0
	511.2363	508.3631	505.3240	0	0	0	0
	507.2411	504.3290	0	0	0	0	0
	503.0313	0	0	0	0	0	0
b=A_{13}	$n_2 \rightarrow$						
$n_1 \downarrow$	524.9927	523.2505	521.0999	518.5961	515.3792	511.9372	506.2728
	522.8557	520.8094	518.4167	515.3031	511.9863	506.4900	0
	520.2148	517.9337	514.9231	511.7325	506.4040	0	0
	517.1472	514.2391	511.1758	506.0150	0	0	0
	513.2511	510.3163	505.3230	0	0	0	0
	509.1541	504.3280	0	0	0	0	0
	503.0303	0	0	0	0	0	0
b=A_{23}	$n_2 \rightarrow$						
$n_1 \downarrow$	523.4091	521.4503	519.1773	516.1619	513.0984	509.8109	506.2728
	521.0733	518.9150	516.0008	513.0881	509.9046	506.4900	0
	518.3498	515.5357	512.7742	509.6944	506.4040	0	0
	514.7666	512.1565	509.1803	506.0150	0	0	0
	511.2353	508.3621	505.3230	0	0	0	0
	507.2401	504.3280	0	0	0	0	0
	503.0303	0	0	0	0	0	0

In Table 6.4, '0' denotes the states that do not exist in the state space. As for this special case $n_3 = 0$, we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion of Theorem 6.2.

6.3 Reserved Channel Schemes

There could be some channels reserved for the 'new calls' in different cells. The handoff calls are always allowed. We only make the decisions when there is a new call coming into a cell. This is different with the Reserved Channel Scheme (RCS) [23,24] which reserve channels for the handoff calls.

6.3.1 Model Formulation

Consider there are U classes of multimedia calls (telephone, video, etc., but for convenience we shall call all of them *calls*) in the network. The other assumptions and notations for this wireless network are as follows.

1. The required bandwidth of class u calls $u = 1, \ldots, U$ is b_u.

2. The network consists of a number (N) of cells.

3. New calls of class u are generated in cell i according to a Poission process with rate λ_{ui}, $i = 1, 2, \ldots, N$. The requested call connection time (RCCT) of a new call at cell i, H_{ui}, is exponentially distributed with means $1/h_{ui}$.

4. The cell residence time, which is defined as the length of time a class u call stays in the cell i and which depends on the velocity and the direction of the mobile terminal, R_{ui}, is exponentially distributed with means $1/r_{ui}$.

5. The probability that a call moves from cell i to a neighboring cell j, given that it moves to a neighboring cell before the call is completed, is $p_{i,j}$, where $\sum_{j=1}^{N} p_{i,j} = 1$. Clearly, $p_{i,i} = 0$ and cell j is a neighboring cell of i if and only if $p_{i,j} > 0$.

6. There are a total of C channels in the network. A number of channels K_{ui} can be reserved for the new calls of class u in cell i if necessary, and the handoff calls among cells are always allowed.

7. Accepting a new call of class u in cell i would contribute R_{ui} units of reward to the system. Let the number of calls in cells be
   $\mathbf{n} = (n_{11}, \ldots, n_{U1}, n_{12}, \ldots, n_{U2}, \cdots, n_{1N}, \ldots, n_{UN})$,
   n_{ui} is the numbers for calls of class u in cell i and the system incurs a holding cost rate $f(\mathbf{n})$ per unit time.

Based on these assumptions, we can build the SMDP model for this wireless cellular network as follows:

1. Let the state variable consists of number of calls in the system, the status of calls leaving, arriving to the system. So $S = \{0, 1, \ldots, \lfloor \frac{C}{b_1} \rfloor\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C}{b_U} \rfloor\} \times \cdots \times \{0, 1, \ldots, \lfloor \frac{C}{b_1} \rfloor\} \times \ldots \times \{0, 1, \ldots, \lfloor \frac{C}{b_U} \rfloor\} \times \{D, A_{11}, \ldots, A_{U1}, A_{12}, \ldots, A_{U2}, \cdots, A_{1N}, \ldots, A_{UN}\}$, where $\lfloor x \rfloor$ denotes the nearest integer to x. In the third set D means a departure from the system, while A_{ui} means an arrival of a new call of class u in cell i. And, $\sum_{u=1}^{U} \sum_{i=1}^{N} n_{ui} b_u \leq C$. In general a state could be written as $\langle \mathbf{n}, b \rangle$, where b stands for the last call event, $b \in \{D, A_{11}, \ldots, A_{U1}, A_{12}, \ldots, A_{U2}, \cdots, A_{1N}, \ldots, A_{UN}\}$.

2. In states $\langle \mathbf{n}, D \rangle$, set a_C as the action to continue, thus $A_{\langle \mathbf{n}, D \rangle} = \{a_C\}$. In states $\langle \mathbf{n}, A_{ui} \rangle$, set a_R as the action to reject the call and a_A as the action to admit, so $A_{\langle \mathbf{n}, A_{ui} \rangle} = \{a_R, a_A\}$.

3. The decision epochs are those time points when a call arriving or leaving the system. For this process, the times between decision epochs are exponentially distributed, and let the distribution of the time between decision epochs starting from state s be

$$F(t|s, a) = 1 - e^{-\beta(s,a)t}, t \geq 0.$$

For each state $s = \langle \mathbf{n}, b \rangle$ and action a, let $\beta_0 = \sum_{u=1}^{U} \sum_{i=1}^{N} (\lambda_{ui} + n_{ui}(r_{ui} + h_{ui}))$, so $\beta(s, a)$ can be written as

$$\beta(s, a) = \begin{cases} \beta_0, & a = a_C, a = a_R, \\ \beta_0 + r_{ui} + h_{ui}, & a = a_A, b = A_{ui}. \end{cases}$$

where $u = (1, 2, \cdots, U), i = (1, 2, \cdots, N)$.

4. Let $q(z|s,a)$ denote the probability that the system occupies state j in the next epoch, if at the current epoch the system is at state s and the decision maker takes action $a \in A_s$. For states $s = \langle \mathbf{n}, b \rangle$, $b = D, a = a_C$ and $b = A_{ui}, a = a_R$, let

   $\mathbf{n}_{ui} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots \max(n_{ui}-1, 0) \ldots, n_{Ui}, \cdots, n_{1N}, \ldots, n_{UN}\}$

   and

   $\mathbf{n}^{ui} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots, n_{ui}+1, \ldots, n_{Ui}, \cdots, n_{1N}, \ldots, n_{UN}\}$,

   $\mathbf{n}_{ui}^{uj} = \{n_{11}, \ldots, n_{U1}, \cdots, n_{1i}, \ldots, n_{ui}-1, \ldots, n_{Ui}, \cdots, n_{1j}, \cdots, n_{uj}+1, \ldots, n_{Uj},$

   $\cdots, n_{1N}, \ldots, n_{UN}\}$,

   the state transition probabilities are

   $$q(z|s,a) = \begin{cases} \lambda_{ui}/\beta_0, & z = \langle \mathbf{n}, A_{ui} \rangle, \\ n_{ui}h_{ui}/\beta_0, & z = \langle \mathbf{n}_{ui}, D \rangle, \\ n_{ui}r_{ui}p_{ij}/\beta_0, & z = \langle \mathbf{n}_{ui}^{uj}, D \rangle. \end{cases}$$

   And, for states $s = \langle \mathbf{n}, b \rangle$, $b = A_{ui}, u = (1, 2, \cdots, U), i = (1, 2, \cdots, N), a = a_A$, since admitting an incoming call migrates the system state immediately, there is

   $$q(z|s, a_A) = q(z|\langle \mathbf{n}^{ui}, D \rangle, a_R).$$

5. Because the system state does not change between decision epochs, the expected discounted reward between epochs satisfies

   $$\begin{aligned} r(s,a) &= k(s,a) + c(s,a)E_s^a\left\{\int_0^t e^{-\alpha\tau} d\tau\right\} \\ &= k(s,a) + c(s,a)E_s^a\left\{[1 - e^{-\alpha t}]/\alpha\right\} \\ &= k(s,a) - \frac{c(s,a)}{\alpha + \beta(s,a)}, \end{aligned}$$

where

$$k(s,a)=\begin{cases}0, & a=a_C, a=a_R,\\ R_{ui}, & a=a_A, b=A_{ui}.\end{cases}$$

and $c(s,a)$ is the holding cost rate if the system is at state s and takes action a. We have the holding cost rate as

$$c(s,a)=\begin{cases}-f(\mathbf{n}), & a=a_C, a=a_R,\\ -f(\mathbf{n}^{ui}), & a=a_A, b=A_{ui}.\end{cases}$$

Based on the assumptions, for the admission control problem, since both the state space S and the action space A_s are finite, the reward function $r(s,a)$ is also finite. From *Theorem 11.3.2* of [13], the optimal policy is a stationary deterministic policy d^∞, so the problem can be reduced to find a deterministic decision rule d. For each deterministic decision rule d, let $q_d(z|s)=q(z|s,d(s))$, $r_d(s)=r(s,d(s))$ and $\beta_d(s)=\beta(s,d(s))$, we have,

$$v_\alpha^{d^\infty}(s)=r_d(s)+\frac{\beta_d(s)}{\alpha+\beta_d(s)}\sum_{z\in S}q_d(z|s)v_\alpha^{d^\infty}(z). \tag{6.5}$$

From equation (6.5), it is seen that if $\beta_d(s)$ is a constant for all state s, the calculation for $v_\alpha^{d^\infty}(s)$ could be simplified. This is the idea of rate uniformization technique. Based on the assumptions, our process fits the condition of *Assumption 11.5.1* of [13], which is $[1-q(s|s,a)]\beta(s,a)\le c, \forall s\in S, a\in A_s$, here c is a constant. So, we can define a uniformization of our process with components denoted by $\sim$. Let $\tilde{S}=S$, $\tilde{A}_s=A_s$, and $c=C*\sum_{i=1}^{N}(\lambda_i+r_i+h_i)$, we have

$$\tilde{q}(z|s,a)=\begin{cases}1-\frac{[1-q(s|s,a)]\beta(s,a)}{c}, & z=s,\\ \frac{q(z|s,a)\beta(s,a)}{c}, & z\neq s.\end{cases}$$

And, for the reward functions, we have

$$\tilde{r}(s,a) \equiv r(s,a)\frac{\alpha+\beta(s,a)}{\alpha+c}.$$

From *Proposition 11.5.1* [13], for each d^{∞} policy and $s \in S$, we have

$$\tilde{v}_{\alpha}^{d^{\infty}}(s) = v_{\alpha}^{d^{\infty}}(s). \tag{6.6}$$

From equation (6.5) and (6.6), the optimality equation of $v(s)$ for maximum $v_{\alpha}^{\pi}(s)$ would have the form of

$$v(s) = \max_{a \in A_s}\left\{\tilde{r}(s,a) + \lambda \sum_{z \in S} \tilde{q}(z|s,a)v(z)\right\}, \tag{6.7}$$

where $\lambda \equiv \frac{c}{c+\alpha}$. After uniformization, the transition process from one state to another can be described by a discrete-time Markov chain which allows fictitious transitions from a state to itself.

From equation (6.7), we have for states with $b = D$,

$$\begin{aligned} v(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha+c}&[-f(\mathbf{n}) + \sum_{u=1}^{U}\sum_{i=1}^{N}\lambda_{ui}v(\langle \mathbf{n}, A_{ui}\rangle) \\ &+\sum_{u=1}^{U}\sum_{i=1}^{N} n_{ui}h_{ui}v(\langle \mathbf{n}_{ui}, D\rangle) + \sum_{u=1}^{U}\sum_{j=1}^{N}\sum_{i=1}^{N} p_{ij}n_{ui}r_{ui}v(\langle \mathbf{n}_{ui}^{uj}, D\rangle) \\ &+(c-\beta_0)v(\langle \mathbf{n}, D\rangle)]. \end{aligned} \tag{6.8}$$

6.3.2 Theoretical Results

Since admitting a call migrates the system state immediately, we have for states $b = A_{ui}$,

$$v(\langle \mathbf{n}, A_{ui}\rangle) = \max\left[v(\langle \mathbf{n}, D\rangle), R_{ui} + v(\langle \mathbf{n}^{ui}, D\rangle)\right].$$

Theorem 6.3 *If $v(\langle \boldsymbol{n}, D\rangle)$ is concave nonincreasing on n_{ui}, $u = 1, \ldots, U, i = 1, \ldots, N$, the optimal policy is a control limit policy.*

Proof: We already know that the optimal policy is a stationary deterministic policy. Let $\triangle v_{ui}(\langle \mathbf{n}, D\rangle) = v(\langle \mathbf{n}^{ui}, D\rangle) - v(\langle \mathbf{n}, D\rangle)$, so $\triangle v_{ui}(\langle \mathbf{n}, D\rangle)$ is ≤ 0 and is nonincreasing, for states with fixed parameters other than n_{ui}, we have the decision rule

$$d(\mathbf{n}) = \begin{cases} a_A, & \triangle v_{ui}(\langle \mathbf{n}, D\rangle) > -R_{ui}, \\ a_R, & \triangle v_{ui}(\langle \mathbf{n}, D\rangle) \leq -R_{ui}. \end{cases}$$

So, if $d(\mathbf{n}) = a_R$, we have $d(\mathbf{n}^{ui}) = a_R$, and so on as u goes through $1, \ldots, U$ and i goes through $1, \ldots, N$. Consequently the optimal policy is a control limit policy (or threshold policy).

Theorem 6.4 *Suppose that the cost function $f(\boldsymbol{n})$ is convex and nondecreasing on $n_{ui}, u = 1, \ldots, U, i = 1, \ldots, N$, and*

$$f(\boldsymbol{n}) + f(\boldsymbol{n}^{uiuj}) \geq f(\boldsymbol{n}^{ui}) + f(\boldsymbol{n}^{uj}),$$

where $\boldsymbol{n}^{uiuj} = \{n_{11}, \ldots, n_{U1}, n_{12}, \ldots, n_{U2}, \cdots, n_{1i}, \ldots n_{ui} + 1 \ldots, n_{Ui}, \cdots, n_{1j}, \ldots n_{uj} + 1 \ldots, n_{Uj}, \cdots, n_{1N}, \ldots, n_{UN}\}$ and $i, j \in 1, \ldots, N$, then (1): $v(\langle \boldsymbol{n}, D\rangle)$ is concave non-increasing on $n_{ui}, i = 1, \ldots, N$, and (2): the optimal policy is a control limit policy.

Proof: We use Value Iteration method to verify it.

1. Set $v^0 = 0$ and substitute this into equation (6.8) to obtain $v^1(\langle \mathbf{n}, D\rangle) = \frac{-f(\mathbf{n})}{\alpha+c}$.

2. Set k=1. We have $v^k(\langle \mathbf{n}, D\rangle)$ which is concave nonincreasing. From Lemma A.1, $v^k(\langle \mathbf{n}, A_{ui}\rangle)$ is concave nonincreasing on $n_{ui}, u = 1, \ldots, U, i = 1, \ldots, N$. And we have,

$$v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{uiuj}, b\rangle) \leq v^k(\langle \mathbf{n}^{ui}, b\rangle) + v^k(\langle \mathbf{n}^{uj}, b\rangle),$$

for $b = A_{ui}, u = 1, \ldots, U, \& i, j \in 1, \ldots, N$.

3. Substituting these v^k back into equation (6.8), we get

$$v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[-f(\mathbf{n}) + \sum_{u=1}^{U}\sum_{i=1}^{N}\lambda_{ui}v^k(\langle \mathbf{n}, A_{ui}\rangle)$$
$$+\sum_{u=1}^{U}\sum_{i=1}^{N}n_{ui}h_{ui}v^k(\langle \mathbf{n}_{ui}, D\rangle) + \sum_{u=1}^{U}\sum_{j=1}^{N}\sum_{i=1}^{N}p_{ij}n_{ui}r_{ui}v^k(\langle \mathbf{n}_{ui}^{uj}, D\rangle)$$
$$+(c - \beta_0)v^k(\langle \mathbf{n}, D\rangle)].$$

so we have for $u = 1, \ldots, U, i = 1, \ldots, N$,

$$v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) - v^{k+1}(\langle \mathbf{n}, D\rangle) = \frac{1}{\alpha + c}[$$
$$\sum_{v=1}^{U}\sum_{w=1}^{N}\lambda_{vw}(v^k(\langle \mathbf{n}^{ui}, A_{vw}\rangle) - v^k(\langle \mathbf{n}, A_{vw}\rangle))$$
$$+\sum_{v=1}^{U}\sum_{w=1}^{N}n_{vw}h_{vw}(v^k(\langle \mathbf{n}_{vw}^{ui}, D\rangle) - v^k(\langle \mathbf{n}_{vw}, D\rangle))$$
$$+\sum_{v=1}^{U}\sum_{j=1}^{N}\sum_{w=1}^{N}p_{wj}n_{vw}r_{vw}(v^k(\langle \mathbf{n}_{vw}^{uivj}, D\rangle) - v^k(\langle \mathbf{n}_{vw}^{vj}, D\rangle))$$
$$+(c - \beta_{ui})(v^k(\langle \mathbf{n}^{ui}, D\rangle) - v^k(\langle \mathbf{n}, D\rangle))$$
$$+(f(\mathbf{n}) - f(\mathbf{n}^{ui}))],$$

where $\beta_i = \beta_0 + r_{ui} + h_{ui}$. Since $v^k(\langle \mathbf{n}, b\rangle)$ are concave nonincreasing on n_i, the combination of these functions $v^{k+1}(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing on n_i.

Also, using equation (6.8),

$$v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uiuj}, D\rangle)$$
$$-v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) - v^{k+1}(\langle \mathbf{n}^{uj}, D\rangle)$$

could be divided into several different parts of $v^k(\langle \mathbf{n}, b\rangle) + v^k(\langle \mathbf{n}^{uiuj}, b\rangle) -$

$v^k(\langle \mathbf{n}^{ui}, b\rangle) - v^k(\langle \mathbf{n}^{uj}, b\rangle)$, where $b = D, A_{ui}$, which are all ≤ 0. So

$$\begin{aligned} & v^{k+1}(\langle \mathbf{n}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uiuj}, D\rangle) \\ \leq\ & v^{k+1}(\langle \mathbf{n}^{ui}, D\rangle) + v^{k+1}(\langle \mathbf{n}^{uj}, D\rangle). \end{aligned}$$

4. Set k=k+1, go back to step 2.

5. As the iteration continues, $v^k(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing, by the *Theorem 11.5.2* of [13], the solution of $v(s)$ is unique, so the value iteration $v^k(s)$ converges to $v(s)$, so $v(\langle \mathbf{n}, D\rangle)$ is concave nonincreasing.

6. From Theorem 6.3, the optimal policy is a control limit policy. The verification is over.

It is seen from Theorem 6.4 that no matter how many channels are reserved, the optimal policy is always a control limit policy if the holding cost functions are convex and nondecreasing.

6.3.3 Simulation Results

Without losing generosity, suppose there are $N = 3$ cells and 2 classes of traffic in this wireless network. The routing probabilities between cells are set in Table 6.5. For simplification, let the total channels be $C = 6$, the cost function be $f(\mathbf{n}) = 2n_{11}^2 + 1.5n_{12}^2 + n_{13}^2 + n_{21}^2 + n_{22}^2 + 0.5n_{23}^2$, the discount factor $\alpha = 0.1$, the bandwidths for two classes are $b_1 = 1, b_2 = 2$, and we set the other parameters for analysis as in Table 6.6 to study the performance of the RCS scheme. We get the actions and $v(s)$ values for states in the Table 6.7 and Table 6.8.

In Table 6.6, λ is the arrival rate of new calls, h is the call connection rate, r is the cell residence rate, R is the reward for new calls, and K is the number of reserved channels.

Table 6.5: Routing Probablities

Cell	1	2	3
1	0	0.4	0.6
2	0.5	0	0.5
3	0.8	0.2	0

Table 6.6: Parameters Setting

Class	Cell	λ	h	r	R	K
1	1	4	6	4	5	0
1	2	2	4	5	4	0
1	3	1	3	2	6	1
2	1	4	6	4	8	0
2	2	2	4	5	7	0
2	3	1	3	2	6	1

Table 6.7 shows the actions for call arrivals when there are no class 2 calls in the system and $n_{13} = 0$, which means there is currently no call of class 1 in cell 3. In Table 6.7, '1' means to admit a new call , and '0' are those states in which the system does not accept any new calls. '-1' denotes those states that do not exist in the state space. As seen from Table 6.7, we have when $n_3 = 0, n_1 = 3$, because of the reserved channel in cell 1, only the arrival of new calls in cell 1 is allowed. Because all the handoffs are allowed in the system, although there is one channel reserved for the new calls in cell 1, the system can still have $n_2 = 4$, this is due to the handoffs from other cells to cell 2. It is seen that when n_1 or n_2 is fixed, there is a threshold value in the decision rule for accepting the new calls in the cells, so the optimal policy is a control limit policy, which is consistent with Theorem 6.4.

In Table 6.8, '-1' denotes the states that do not exist in the state space. As for this special case $n_3 = 0$, we can see that the values of $\triangle v(s)$ are nonincreasing in both $n_1 \downarrow$ and $n_2 \rightarrow$ directions, which fits our theoretical conclusion of Theorem 6.4.

Table 6.7: Actions for calls arrival in Cell 3

b = A_{13}	n_{12}					
n_{11}	1	1	1	1	0	0	0
	1	1	1	1	0	0	-1
	1	1	1	1	0	-1	-1
	1	1	1	0	-1	-1	-1
	0	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1
b = A_{23}	n_{12}						
n_{11}	1	1	1	1	1	0	0
	1	1	1	0	0	0	-1
	1	1	0	0	0	-1	-1
	1	0	0	0	-1	-1	-1
	0	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1

Table 6.8: $v(s)$ for states

b=D	$n_2 \rightarrow$						
$n_1 \downarrow$	495.6301	494.3587	492.7504	490.6656	487.9983	484.6509	481.3739
	494.0583	492.5614	490.6673	488.2919	484.9824	481.7605	0
	492.0341	490.2527	488.0435	485.033	481.8224	0	0
	489.3254	487.3064	484.458	481.4742	0	0	0
	486.2132	483.4791	480.7432	0	0	0	0
	482.243	479.6811	0	0	0	0	0
	478.3212	0	0	0	0	0	0

6.4 Numerical Analysis

As seen from the previous sections, to compare the NPS and RCS schemes, we set the same parameters and routing probabilities among the cells, the only difference is on the number of channels reserved.

Compare the actions in Table 6.3 and Table 6.7, based on the reward, the actions with NPS scheme is to accept any incoming arrival calls. But with the RCS scheme, due to reserved channel for the new call in cell 3, the arrival new calls in cell 1 and 2 can not be always accepted.

We have shown the actions of $n_3 = 0$, using the same parameter setting, We show the actions and $v(s)$ values for states in the when $n_1 = 0$ under NPS and RCS schemes.

Table 6.9: Actions for calls arrival in Cell 1, NPS

b = A_{11}	n_{12}						
n_{11}	1	1	1	1	0	0	0
	1	1	1	1	0	0	-1
	1	1	1	1	0	-1	-1
	1	1	1	0	-1	-1	-1
	0	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1

Table 6.10: Actions for calls arrival in Cell 1, RCS

b = A_{11}	n_{12}						
n_{11}	1	1	1	1	0	0	0
	1	1	1	1	0	0	-1
	1	1	1	1	0	-1	-1
	0	0	0	0	-1	-1	-1
	0	0	0	-1	-1	-1	-1
	0	0	-1	-1	-1	-1	-1
	0	-1	-1	-1	-1	-1	-1

Compare the actions in Table 6.9 and Table 6.10, due to the channel reserved for the A_{13} and A_{23}, the new calls of class 1 is not always allowed.

The simulation and the above analysis has revealed that:

1. Regardless of the numbers of channels reserved, the optimal policy is a control limit policy or threshold policy if the cost functions follow some special properties. But the numbers of reserved channels do affect the threshold values in the policy.

2. The values of total expected discounted reward $v(s)$ are different if the numbers of reserved channels are changed.

3. When the reward of a class call is fixed, simply increasing channel capacities does not change the optimal policy for this class calls.

In this chapter, we consider multimedia traffic in a wireless network with multiple cells. There are routing probabilities between cells. Assume that the arrival process of each new class calls in each cell is a Poisson process, the call connection time and cell residence time follow exponential distributions, accepting each new call would contribute some units of reward to the system and the system incurs a holding cost per unit time for the calls in the system. We find that if the holding cost function is convex nondecreasing, the optimal policy call admission policy for calls is a control limit policy. The numerical analysis has proven our theoretical conclusion and described the relationship between parameters. The result of this chapter could easily be used in designing wireless networks for optimal performance.

Chapter 7

Conclusion and Suggestions for Future Research

This dissertation has investigated the optimal call admission control (CAC) policies for Non Priority Scheme (NPS) and Reserved Channel Scheme (RCS) in wireless cellular networks. This dissertation focused on the optimization problem of when to admit or reject a call in order to achieve the maximum reward. By establishing an infinite horizon discounted Semi Markov Decision Process (SMDP) model, this dissertation has shown that the optimal policies are state-related control policies for NPS as well as RCS. The numerical analysis has proven our theoretical conclusion and established the relationship between system parameters. The result of this dissertation could easily be used in designing wireless networks for optimal performance.

This work can be extended to include the following:

1. General Distribution for new calls and handoff calls

2. Handling handoff calls in the network

3. Relationship between control limit and parameters

Appendix A

Definitions and Lemmas

We make definitions in the below. This definitions are for the cost functions that are used in models.

Definition: In this dissertation we say a discrete function $f(i)$ is convex nondecreasing if $f(i+1) - f(i) \geq f(i) - f(i-1) \geq 0, \forall i$, or concave nonincreasing if $f(i+1) - f(i) \leq f(i) - f(i-1) \leq 0, \forall i$.

Definition: In this dissertation we say a two dimensional discrete function $f(i,j)$ is convex nondecreasing on i if $f(i+1,j) - f(i,j) \geq f(i,j) - f(i-1,j) \geq 0$, or concave nonincreasing on i if $f(i+1,j) - f(i,j) \leq f(i,j) - f(i-1,j) \leq 0$.

Definition: In this dissertation we say a function $f(n), n = (n_1, \ldots, n_I)$ is convex nondecreasing on n_i if $f(n^i) - f(n) \geq f(n) - f(\bar{n}_i) \geq 0$, or concave nonincreasing on n_i if $f(n^i) - f(n) \leq f(n) - f(\bar{n}_i) \leq 0$.

The following lemmas A.1 and A.2 are from [25].

Lemma A.1 *Suppose that $h(i)$ is concave nonincreasing, then*

$$g(i) \equiv \max\{h(i), R + h(i+1)\},$$

is concave nonincreasing.

Lemma A.2 *Suppose that $f(i,j)$ is convex and nondecreasing on both i and j, and*

$$f(i,j) + f(i+1,j+1) \geq f(i+1,j) + f(i,j+1),$$

then for any $R \geq 0$,

$$g(i,j) + g(i+1,j+1) \leq g(i+1,j) + g(i,j+1),$$

where

$$g(i,j) \equiv \max\{-f(i,j), R - f(i+1,j)\}.$$

Based on the lemmas A.1 and A.2, we have

Lemma A.3 *Suppose that $g_n(i), n = 1, 2, \ldots$ is concave nonincreasing, and let $c_n \geq 0, n = 1, 2, \ldots$ be constants, then*

$$g(i) = \sum_n \{c_n * g_n(i)\},$$

is also concave nonincreasing.

Lemma A.4 *Suppose that $f(i,j)$ is convex and nondecreasing on both i and j, and*

$$f(i,j) + f(i+1,j+1) \geq f(i+1,j) + f(i,j+1),$$

then for any $R \geq 0$*,*

$$g_1(i,j) \equiv \max\{-f(i,j), R - f(i+1,j)\},$$
$$g_2(i,j) \equiv \max\{-f(i,j), R - f(i,j+1)\},$$

are concave nonincreasing on both i *and* j*.*

Bibliography

[1] A. Kumar, D. Manjunath, J. Kuri, *Communication Networking: An Analytical Approach* . Elsevier Inc, 2004, ISBN 0-12-428751-4.

[2] A. Z. Melikov, and A. T. Babayev, *Refined approximations for performance analysis and optimization of queueing model with guard channels for handovers in cellular networks* . Computer Communications, vol. 29, no. 9, pp. 1386-1392, 2006.

[3] B. Moretta, I. Zideins, *Admission controls for Erlang's loss system with service times distributed as a finite sum of exponential random variables.* Applied Mathematics & Decision Sciences, 2(2), 119-132 (1998).

[4] C. J. Ho, C. T. Lea, *Improving call admission policies in wireless networks.* Wireless Networks, 5, 1999 (257-265).

[5] D. P. Gaver, P. A. Jacobs and G. Latouche, *Finite birth-and death models in randomly changing environments* . Advances in Applied Probability, vol.16, 715-731, 1984.

[6] D. Hong and S. S. Rappaport, *Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures* . IEEE Transactions on Vehicular Technology, vt-35, no.3, pp.77-92, 1986.

[7] H. Chen, L. Huang, S. Kumar and C. Jay Kuo, *Radio Resource Management For Multimedia QoS Support In Wireless Networks.* John Wiley & Sons, 1983, ISBN 0-471-09942-2.

[8] J. Medhi, *Stochastic Models In Queuing Theory*, 2nd ed. Academic Press, 2003, ISBN 0-12-487462-2.

[9] J. F. Hayes, T. Babu, *Modeling and Analysis of Telecommunications Networks.* Wiley-Interscience, 2003, ISBN 0-471-34845-7.

[10] J. Walrand, P. Varaiya, *High Performance Communication Networks* , 2nd ed. Morgan Kaufmann Publishers, 2000, ISBN 1-55860-574-6.

[11] K. W. Ross, *Multiservice Loss Models for Broadband Telecommunication Networks.* Springer-Verlag London Limited, 1995, ISBN 3-540-19918-7.

[12] M.F. Neuts, *Matrix-geometric solutions in stochastic models* . The John Hopkins University Press, 1981.

[13] M. L. Puterman, *Markov Decision Process: Discrete Stochastic Dynamic Programming.* Wiley-Interscience, 2005, ISBN 0-471-72782-2.

[14] N. Bartolini, I. Chlamtac, *Call Admission Control In Wireless Multimedia Networks.* IEEE PIMRC, 2002.

[15] P. V. Orlik and S. S. Rappaport, *A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions* . IEEE Journal on Selected Areas in Communications, vol.16, no.5, pp. 788-803, 1998.

[16] S. Lippman, *Applying a New Device in the Optimization of Exponential Queuing Systems.* Operations Research, 23:687-710, 1975.

[17] S. M. Ross, *Stochastic Process.* John Wiley & Sons, 1983, ISBN 0-471-09942-2.

[18] S. S. Rappaport, *The multiple-call hand-off problem in high-capacity cellular communications systems* . IEEE Transactions on Vehicular Technology, vol.40, no.3, pp.546-557, 1991.

[19] Toni Janevshi, *Traffic analysis and design of wireless IP networks* . Artech House, Inc., 2003.

[20] W. Li and A. S. Alfa, *A PCS network with correlated arrival process and splitted-rate channels* . IEEE Journal on Selected Areas in Communications, vol. 17, no.7, 1318-1325, 1999.

[21] W. Li and A. S. Alfa, *Channel reservation for hand-off calls in a PCS network* . IEEE Transactions on Vehicular Technology, vol. 49, no. 1, pp. 95-104, 2000.

[22] W. Li, X. Chao, *Call Admission Control for an Adaptive Heterogeneous Multimedia Mobile Network.* IEEE Transactions On Wireless Communications, Vol.6, No.2, Feb 2007.

[23] W. Li, Y. Fang, *Performance Evaluation of Wireless Cellular Networks with Mixed Channel Holding Times.* IEEE Transactions On Wireless Communications, In Press.

[24] W. Li, Y. Fang, R. Henry, *Actual Call Connection Time Characterization for Wireless Mobile Networks Under a General Channel Allocation Scheme.* IEEE Transactions On Wireless Communications, Vol. 1, No. 4, October 2002.

[25] X. Chao, H. Chen, W. Li, *Optimal Control for a Tandem Network of Queues with Blocking.* ACTA Mathematicae Applicatae Sinica, Vol.13, No.4, Oct 1997.

[26] Y. Fang and Y. Zhang, *Call admission control scheme and performance analysis*

in wireless mobile networks . IEEE Transaction on Vehicular Technology, vol. 51, no.2, pp.371-382, 2002.

[27] Y. Fang, *Thinning scheme for call admission control in wireless networks* . IEEE Transaction on Computers, vol. 52, no.5, pp.685-687, 2003.

[28] Y. Fang and I. Chlamtac, *Teletraffic analysis and mobility modeling of PCS networks* . IEEE Transactions on Communications, vol.47, no.7, pp.1062-1072, 1999.

[29] Y. Fang, I. Chlamtac and Y. B. Lin, *Channel occupancy times and handoff rate for mobile computing and PCS networks* . IEEE Transactions on Computers, vol.47, no.6, pp.679-692,1998.

[30] Y. B. Lin, S. Mohan and A. Noerpel, *Queueing priority channel assignment strategies for PCS handoff and initial access* . IEEE Transactions on Vehicular Technology, vol.43, no.3, pp.704-712, 1994.

[31] Y. B. Lin, S. Mohan and A. Noerpel, *PCS channel assignment strategies for hand-off and initial access* . IEEE Personal Communications, vol.1, no.3, pp.47-56, 1994.

[32] Y. Kyung, S. Shaler, *Optimal Service-Rate control of M/G/1 Queuing Systems Using Phase Methods.* Applied Probability, 15:616-637, 1983.