BOSTON UNIVERSITY
COLLEGE OF ENGINEERING

DISSERTATION

PERSONAL COMMUNICATIONS SERVICES (PCS) NETWORKS: MODELING AND PERFORMANCE ANALYSIS

by

YUGUANG FANG

M.S., Qufu Normal University, PRC, 1987
Ph.D, Case Western Reserve University, 1994

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
1997

Approved by

First Reader

Dr. Thomas G. Kincaid, Professor and Associate Dean

Department of Electrical and Computer Engineering

Boston University

Second Reader

Dr. Imrich Chlamtac, Professor

Distinguished Chair of Telecommunications

Erik Josson School of Engineering and Computer Science

The University of Texas at Dallas

Third Reader

Dr. Lev Levitin, Distinguished Professor

Department of Electrical and Computer Engineering

Boston University

Fourth Reader

Dr. David Castañon, Associate Professor

Department of Electrical and Computer Engineering

Boston University

Dedicated

to

My Beloved Parents

and

My Beloved Parent-in-Laws

ACKNOWLEDGMENT

I am greatly indebted to my advisors, Thomas G. Kincaid and Imrich Chlamtac for their great inspiration, excellent guidance, deep thoughts and friendship during the course of my graduate studies.

Thanks are due to my dissertation committee: Professor Lev Levitin and Professor David Castañon for their helpful discussions.

I am grateful to Dr. Imrich Chlamtac, Professor and Distinguished Chair of Telecommunications at the University of Texas at Dallas, and Dr. Yi-Bing "Jason" Lin, Professor at National Chiao Tung University, Taiwan, who were willing to provide the research topic and unselfishly contribute their insightful ideas to the research.

Special thanks are due to the Kincaid family: Betsy and Tom, who not only constantly encouraged me in many ways, but also shared their view of life with me and my family. It is a great pleasure to know them. Special thanks are also extended to the Cotter family: Evelyn and Sarah, for their friendship.

I also express my appreciation to all the faculty, staff and my fellow students in the Department of Electrical and Computer Engineering. In particular, I extend my thanks to Ms. Loraine Walsh for her constant help, my office mates Mr. Chester Ruszczyk and Ms. Hongbiao Zhang for their helpful discussions, and my colleagues Ms. Tong Liu, Mr. Jason Redi and Ms. Chiara Petrioli in the Advanced Communications Laboratory for their help in many ways.

It is always an inspiration whenever I think of my high school teachers Mr. Zhaoxiang Jiang and Mr. Wen-Heng Xu, who played such important roles in my early education.

Finally, I would like to thank my dear wife Jennifer Yih-Ling Lu and my son Andrew Ray Fang for their love, patience, constant encouragement and understanding, "Guys, this is for you!"

PERSONAL COMMUNICATIONS SERVICES (PCS) NETWORKS: MODELING AND PERFORMANCE ANALYSIS

(Order No.)

YUGUANG FANG

Boston University, College of Engineering, 1997
Advisor: Thomas G. Kincaid, Professor,
Electrical and Computer Engineering

Co-advisor: Imrich Chlamtac, Professor,
The University of Texas at Dallas

Abstract

In the classical wireline and cellular networks, it is often assumed that the call holding times, cell residence times and channel occupancy times are exponentially distributed. However, as the cell sizes in PCS networks become smaller and the services provided over such networks become more integrated, the call holding times and cell residence times are no longer exponentially distributed. Hence the analytical results of classical traffic theory are not applicable to PCS network performance evaluation.

In this research, we study the following performance metrics for PCS networks under general distributions of call holding times and cell residence times: cell traffic, channel occupancy times, handoff rate, call dropping probability (call incompletion probability), call completion probability, and effective call holding times. We develop new analytical approaches to tackle this general model. We give analytical characterization of cell traffic and obtain necessary and sufficient conditions for channel occupancy times to be exponentially distributed and for the cell traffic to be Poissonian. Analytical results are obtained for handoff rate, call dropping probability and call completion probability, distributions of channel occupancy times, the distributions of effective call holding times, and the expected effective call holding times for a complete call and for an incomplete call. We show how the expected effective call holding times for a complete call and for an incomplete call can be used to develop new billing rating plans for PCS networks. Three strategies of billing record updating for billing system monitoring and fraud detection are proposed and analyzed. Our results indicate that our new approaches can tackle situations that the classical traffic theory can not handle analytically. The new analytical tools we develop in this research open a new avenue in modeling and performance analysis for the emerging PCS networks.

Contents

List of Figures

Chapter 1

INTRODUCTION

Wireless Personal Communications has captured the intensive attention of both academia and industry, and the Personal Communications Services (PCS) networks is an integral part of this area ([6]). PCS networks allow users communicate as they move or roam across different service providers' domains and to provide seamless communications services including voice, data and multimedia at any time and at any place. These networks allow the use of small low-powered lightweight portables with radio interface to base stations in microcells of typically less than 1 km diameter ([34, 19, 25, 54]). PCS networks is one of the fastest growing and largest revenue-generating areas in telecommunication industry. Intensive research, development and standardization activity on a range of related aspects are being carried out at both the global and regional levels. The design and analysis of such PCS networks are more complex than the classical wired networks such as POTS (Plain Old Telephone Systems). Analytical results in classical traffic theory are not applicable because of the mobility of users and the complex traffic of new emerging integrated services. Hence, traffic engineering for networks supporting mobile and personal communications services has reached a new dimension in teletraffic theory and needs to be carefully investigated ([33]).

In this research, we focus on the modeling and performance analysis of PCS networks. Before we discuss our main topics, we give a brief overview of wireless personal communications on which the PCS networks are based.

1.1 A Brief History

The first wireless telephone service, which was used to interconnect mobile users (typically in automobiles) to the public wired telephone networks, was introduced in the United States in 1946 by AT&T. Frequency modulation (FM) transmission was used between the fixed stations and mobile users. The quality of service was not good due to transmission noise and multipath fading. However, the demand for mobile services was on the rise, hence these systems were not efficient anymore. In the mid-1960s, the Bell System introduced the Improved Mobile Telephone Services (IMTS), which markedly improved the systems.

In the late 1960s and early 1970s, the cellular concept was conceived and was then used to improve the system capacity and frequency efficiency. The idea behind "cellular" is to divide the service area into many small regions (cells) according to the transmission power constraints and availability of spectrum. In these cellular networks, the frequency used in one cell can be reused in another cell which is located far enough to avoid co-channel interference.

The late 1970's and early 1980's saw the popularity of cellular services increase. In North America, the AMPS (Advanced Mobile Phone Systems) and the narrow-band AMPS were developed. Similar systems were also developed in other countries: TACS (Total Access Communications Systems) in the United Kingdom, NMT (Nordic Mobile Telephone) in Scandinavia, C450 in West Germany.

With the development of digital technologies and microprocessing computing power in the late 1980's and up to today, enormous interest emerged in digital cellular systems, which promised higher capacity and higher quality of services at reduced cost. In these systems, all signals, including voice and control signals, are digitized for transmission. GSM (Global Systems for Mobile Communications) was developed in Europe, IS-54/IS-95 were developed in the United States and JDC (Japanese Digital Cellular) was developed in Japan. In 1993, a digital system was placed in service in some parts of the United States. GSM became the standard in Europe and some other countries, while the United States is still making a compromise between the analog cellular systems (AMPS/N-AMPS) and the digital cellular systems (IS-54/IS-95).

As the data services such as Internet service became much more in demand, the integration of various services such as data, voice and multimedia is attracting tremendous interest. In early 1990, personal communications services (PCS) was conceived and investigation began. PCS is likely to explode worldwide since the market demand is still rising.

Historically, wireless personal communications has undertaken three stages or generations, which can be shown in Figure 1.1 Analog cellular systems belong to the first generation; the major service provided is voice. Second generation cellular systems use digital technologies to provide better quality of service including voice and limited data with higher system capacity and lower cost. The third generation is targeted to the future where integrated services can be supported; PCS networks belong to this category. Most of the services supported on the third generation cellular systems are wireless extensions of the Integrated Services Digital Network (ISDN). WACS (Wireless Access Communications System) and PACS (Personal Access Communications System) are two major standard systems in PCS networks.

1.2 What is PCS?

As Cox ([6]) pointed out, there is no clear definition for PCS. He said that "it is not one technology, not one system, and not one service, but encompasses

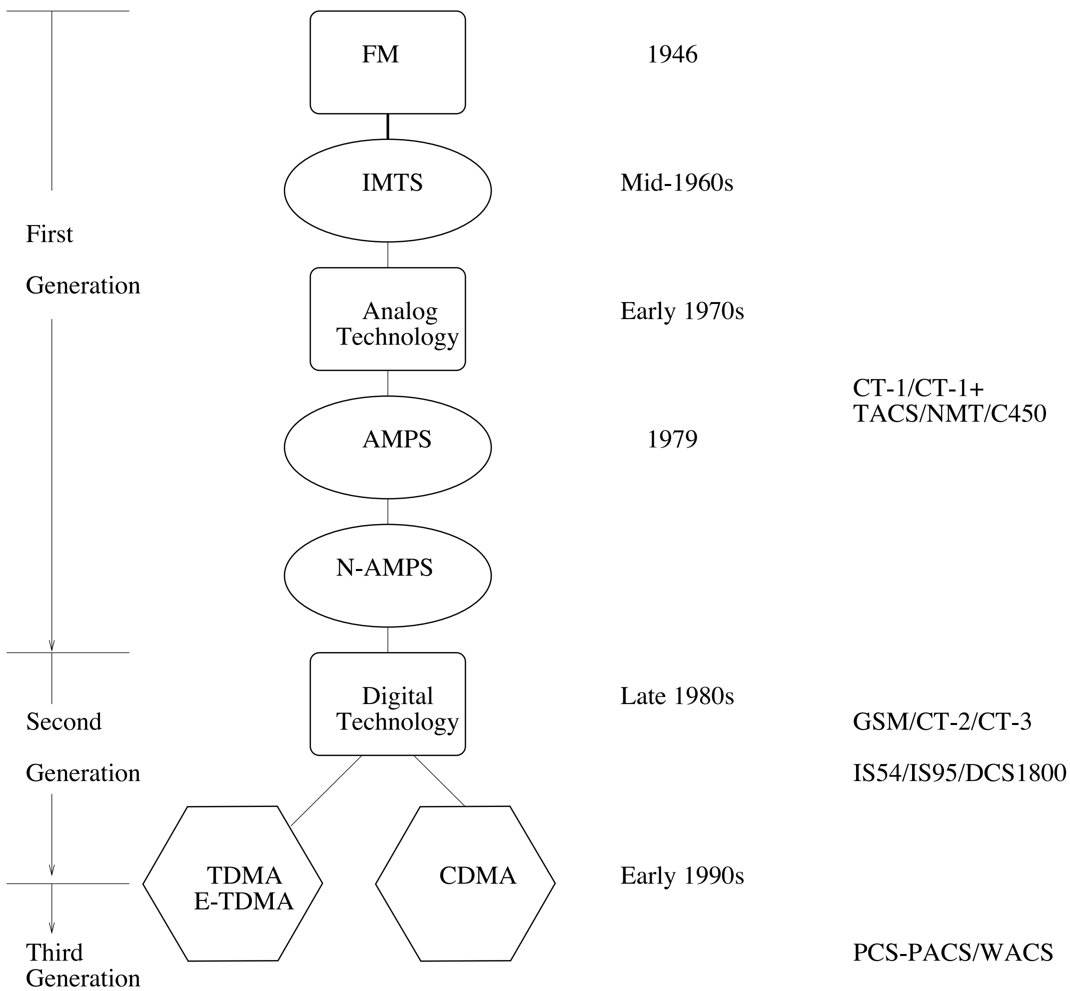


Figure 1.1: Evolution of wireless communications

many technologies, systems and services optimized for different applications". It also has different names at different places with different research groups, such as PCN (Personal Communications Network), UMTS (Universal Mobile Telecommunication System) and FPLMTS (Future Public Land Mobile Telecommunication Systems). However, non-rigorous descriptive definitions can be used to understand the essence of the PCS systems or networks. PCS is often distinguished from cellular telephony by the promise to provide any type of service to anyone, anywhere and at any time with ubiquitous coverage. It developed from the classical cellular telephone systems by the following two approaches: either increasing the cell size to increase the capacity and widespread coverage (like cordless phones or satellite mobile systems) or decreasing the cell size (microcells) and power levels (like vehicular cellular radios and hand-held portables). Both approaches appear to be aimed toward a more widespread convenient form of wireless personal communications. It has been recognized ([53]) that PCS standards fall naturally into two categories: "high tier", supporting macrocells and high-speed mobility; and "low tier", optimized for low power and low complexity. The requirements for PCS networks (as in UMTS from RACE projects in Europe ([57]) or in PACS from Bellcore) can be summarized as follows:

- Spectrally efficient communication system by which every user can exchange information with anyone, anywhere and at any time;
- High capacity to support high density service usage;
- Supports wide a variety of services;
- Indoor coverage (cordless telephone, wireless local area networks (WLAN)) and/or outdoor coverage (cellular radios, WLAN, wide area wireless data, radio paging, extended cordless telephone, satellite-based wireless);
- Light, pocket-sized, low cost handset/portable with low power consumption;
- High service quality;

- Access to different services (two-way voice, data, messaging and video) possible from single terminal type;
- Communication security (encryption);
- Fair charging (flat rate planning).

The most important thing in PCS networks is that they will support three types of mobility: ([21])

- *personal mobility*—enabling relocatable association between user subscription and actual terminal used;
- *terminal mobility*—maintaining an established connection under both static and dynamic conditions, that is, with both fixed and mobile terminal;
- *service mobility*—enabling management of user service profiles.

Thus, a user on the move can, if he/she likes, always communicate with anybody or any type of services anywhere at any place. The proposed services for the PCS networks ([19]) include telephony, teleconference, voice mail, program sound, video telephony, video conference, remote terminal, user profile editing, telefax (Group 4), voiceband data, database access, message broadcast, unrestricted digital information, navigation, location and others which remain to be defined. Service providers will require higher-complexity protocols in the physical link layer because of the unpredictable nature of the radio propagation environment and the inherent mobility in a PCS network. PCS networks will concentrate on the service quality, system capacity, and personal, terminal and service mobility issues.

1.3 Performance Issues of PCS Networks

As in classical cellular telephony, the service area in a PCS network is also divided into cells (cells/microcells/picocells) according to the carrier-to-interference (C/I) requirements, geography and the users' terminal density. Each cell is equipped with a base station and with a number of radio channels assigned according to the available frequency spectrum and frequency reuse pattern. Any portable residing in a cell can communicate through a radio link with the base station located in the cell, which communicates with the Mobile Switching Center (MSC), which is in turn connected to the Public Switching Telephone Networks (PSTN). When a user initiates or receives a call, he/she may roam around in the coverage of the PCS network. If the portable user moves from one cell to another, and the call from/to the portable has not finished, the network has to handoff the call from one cell to another at the cell boundary crossing without user's awareness of handoff and without much degradation of the service quality. Because the number of channels available in a cell is limited, when a portable makes a call, the new originating call may be blocked. Similarly, when the unfinished call moves from one cell to another, the handoff may not be successful, hence the handoff call is forced to terminate.

The major difference between the classical POTS networks and PCS networks is the mobility of users and the related call handling procedure. In POTS, although users may experience busy tones when a call is made, when connected the call will continue without any interruptions until finished. However, in PCS networks (including classical cellular networks), users not only experience busy tones, but also face the forced termination during the handoff process. In particular, in order to increase the system capacity, the cell sizes of PCS networks are reduced, and handoff effects increase. Since users are much more sensitive to handoff blocking than to new call blocking, the PCS service providers have to pay much more attention to the handoff process in the cell channel allocation ([11, 29, 56]). Hence the analysis for

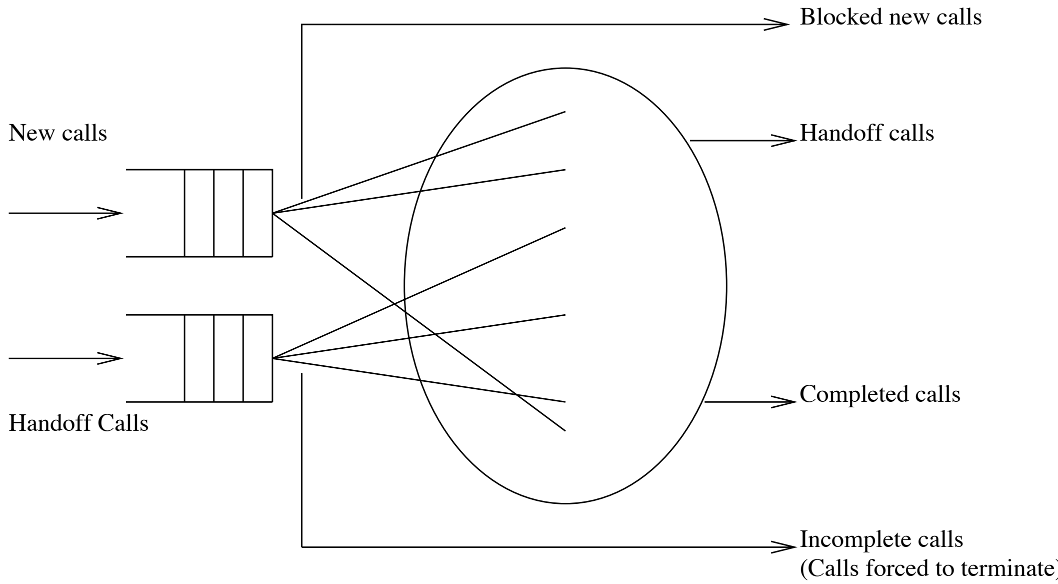


Figure 1.2: New and handoff traffic processes in a typical cell

PCS networks become more complicated. Figure 1.2 shows the traffic processes in a typical cell. New calls and handoff calls arrive at the cell. Based on the availability of channels in the cell, new calls can be blocked or put through and arriving handoff calls can be dropped out or be handed off successfully to the cell. The departed calls from the cell either complete or are handed off to other cells. The following performance metrics may be used to evaluate the PCS network performance:

- *Call blocking probability*—the probability that a new call is blocked.
- *Handoff blocking probability*—the probability that a handoff call is blocked.
- *Handoff Rate*—The average number of handoffs per call.
- *Call dropping probability*—the probability that a call terminates due to a handoff failure. This probability can be easily computed from the call completion probability.
- *Call completion probability*—the probability that a call completes without any blocking and handoff failures.

- *Effective call holding time for a complete call*—the time duration from the call initiation to the finish without any blocking, i.e., the time the complete call spends in the PCS networks.

- *Effective call holding time for an incomplete call*—the time duration from the call initiation to the instant the call is forced to terminate, i.e., the time the incomplete call spends in the PCS networks.

- *Channel occupancy time for a new call*—the time a new call occupies a channel in a cell.

- *Channel occupancy time for a handoff call*—the time a handoff call occupies a channel in a cell.

- *Channel occupancy time*—the time any call (either a new call or a handoff call) occupies a channel in a cell.

- *Handoff call traffic rate*—the handoff call arrival rate to a cell.

Call blocking probability and handoff blocking probability are two measures which are commonly used for the PCS network design and performance evaluation. They are usually the design parameters specified in practical design. Everitt ([11]) emphasized the call dropping probability from the perception of users. This research focus on the call dropping probability (equivalently, the call completion probability) and the rest of the above important quantities under very general assumptions.

1.4 Literature Survey

Intensive research activity for the call blocking probability and handoff blocking probability is reported in the current literature. Hong and Rappaport ([24]) systematically studied the traffic model and performance for cellular systems and obtained some results for call blocking probabilities for new calls and handoff calls using $M/M/m$ queueing system theory. Xie and Kuek ([69]) presented an approximate analysis for Hong-Rappaport model and made some very interesting observations. Rappaport and his coworkers ([59, 60, 58, 61]) also considered the performance of cellular systems with different services and mixed platforms, and the analytical tools may be generalized to the performance of PCS networks with more general traffic conditions. Tekinary and Jabbari ([66]) proposed an analytic model for handoff prioritization scheme, and the $M/M/c/c$ analysis was carried out. Lin et al ([40]) proposed a more general model for channel assignment strategies for PCS network handoffs and obtained some interesting results when some of the traffic conditions are relaxed. Foschini et al ([18]) presented an approximation using a simple *ad hoc* Erlang-B formula based on an equivalent traffic load. Del Re et al ([8]) proposed a new dynamic channel allocation technique, and using the Erlang-B formula they found an analytical expression for the blocking probability. Yoon and Un ([71]) compared three call handling schemes and their performances, assuming the arrival traffic is Poissonian and the departure process is Poissonian. Yum and Yeung ([72]) studied the cellular systems with directed retry and the performance showed improvement. Nanda ([50]) developed some experimental models for teletraffic and obtained a result for handoff rate. Pollini ([56]) summarized most handoff strategies for PCS networks and pointed out the future research trends in that area. Katzela and Naghshineh ([29]) presented an excellent survey of channel allocations for PCS networks.

However, for most of models and systems described in the literature cited above, it is assumed that the cell traffic (the merged traffic from the new calls and handoff calls) is Poissonian and channel occupancy time is exponentially distributed, thus the blocking probabilities can be obtained by the famous Erlang-B formula or the open queueing network theory. Under the same assumption, the handoff probability and call dropping probability can be derived ([34, 50]). Although these assumptions may be valid for certain simple PCS networks, they may not be true for most PCS networks which support the new emerging integrated services. Guerin ([20]) observed that the channel occupancy time may not be exponentially distributed. Bolotin ([4]) found that even the call holding times for certain networks such as the Common-Channel Signaling (CCS) networks are not exponentially distributed, hence they will not be exponentially distributed for PCS networks which consist of various services.

In order to analyze PCS networks, we propose the following two-level model. The first level of the model uses the number of radio channels in cells as an input parameter to determine the new call blocking probability and the handoff call blocking probability. The second level of the model uses the new call blocking probability and handoff call blocking probability as input parameters to study the call completion probability and the effective call holding times.

In this research, we mainly focus on the second level of the model. The importance of the second level is that the PCS service providers may use the expected call holding times for both complete calls and incomplete calls to design appropriate flat rate billing programs. It is not fair to charge the same rate for the customer who is forced to terminate as for those who complete their calls. Terminated calls should be charged a lower rate. However, it is not fair either for the service providers to use this lower rate for the incomplete calls which have already used the service for a long time. Therefore, this research will be very significant for both PCS service providers and customers.

1.5 PCS Network Modeling and Assumptions

A good model in mathematical modeling should satisfy two criteria: it should be general enough to capture the essence of the problem, while at the same time it is analytically tractable. In this section, we present a model for the PCS networks and a set of reasonable assumptions which enable us to give analytical results for the performance of PCS networks. In this section, we also give definitions used in the subsequent development.

When a mobile user attempts to make a call connection to a PCS network, depending on the channel availability, the call may be blocked or connected. The duration of the requested call connection is referred to as the *call holding time* (for example, in wired telephone networks call holding time is the time from the start of the successful connection to the finish of the call). When the call is connected, the call may be complete after some successful handoffs or may be incomplete because of a failed handoff. The probability that a call is complete after the call is connected is called the *call completion probability*, the probability that a call is not blocked but forced to terminate after a few successful handoffs is called *call dropping probability*. It is directly obtained from the call completion probability (e.g., one minus the call blocking probability and minus the call completion probability). We therefore concentrate on the call completion probability in this research. We shall call the duration of an actual call connection for an incomplete call the *effective call holding time of an incomplete call*, while the duration of a call connection of a complete call is called the *effective call holding time of a complete call*. In the wired POTS, these three quantities are the same because there is no failure during a call. However, for PCS networks, these quantities are different and useful for the network performance and billing rating programs. The time that a mobile user spends in a cell is called the *cell residence time*. The cell residence times $t_2, t_3, \cdots$ are shown in Figure 1.3. The effective call holding times and the call completion probability depend on the call holding times, the cell residence times, and the PCS network (new or handoff) call blocking probabilities.

Two main approaches can be identified in the literature for modeling call holding times and cell residence times. One uses the geographic cell (hexagonal) shapes and distribution assumption of the mobiles' speeds, distances and directions to determine the distribution of cell residence times and call holding times ([24, 50, 68]). This approach faces a difficulty when applied to existing cellular systems since in reality

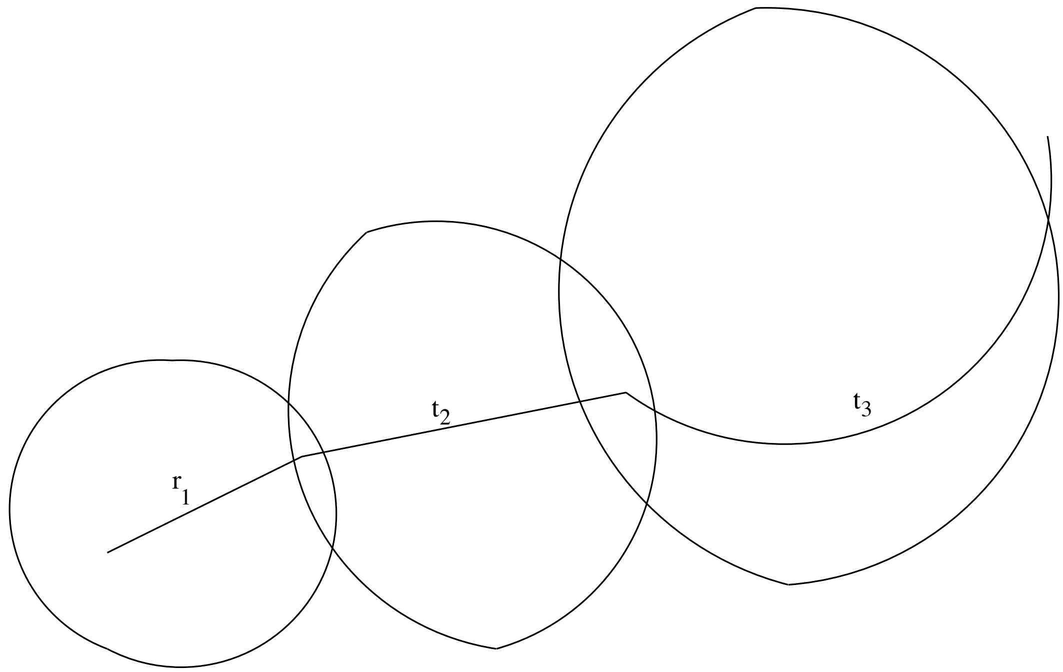


Figure 1.3: The cell residence times $t_2, t_3, \cdots$

cell shapes are often highly irregular, the speeds of mobiles and distances covered by the mobiles are highly random (considering a highly populated area), and the directions of mobiles may vary in zigzag fashion. It becomes, therefore, very hard to characterize analytically the cell residence times using these modeling assumptions.

The second approach treats the cell residence times and call holding times directly using information measured from the PCS field trials. Distribution models, such as exponential distribution and Gamma distribution, have been used to approximate the distributions of call holding times and cell residence times using data from field tests (see [40, 43] and references therein). It is well-known that the exponential distribution can be used for a one parameter approximation of the measured data, while Gamma distribution can be used for secondary approximation ([27]). It is also known ([30]) that mixed Erlang distribution (hyper-Erlang distribution in this research) can be used to approximate any specific non-lattice distributions. Hence the application of these distributions to model the call holding times and cell residence times in the emerging PCS networks appears to be more practical when field data are available. We will take the second approach in this research.

In the traditional wired-line telephone models, the call holding times are usually assumed to be exponentially distributed, which has been shown to be a reasonable approximation for measured data. This assumption has been used in the past for PCS network analysis for the sake of tractability ([20, 24, 68, 72]). When the call holding times are exponentially distributed, Lin et al ([40, 43]), Rappaport et al ([24, 60, 61]), Yum and Yeung ([72]) and Tekinary and Jabbari ([66]) studied the performance of channel assignment strategies and obtained some analytical results for the forced termination probability and the new call blocking probability.

As emerging PCS networks are poised to provide various new integrated services and are expected to attract more users while changing the users' calling habits, the call holding times are not be expected to have the exponential (Erlang) distribution, as already observed by Guerin ([20]) and Bolotin ([4]). More general distributions for call holding times are needed to model the emerging PCS networks. Here we propose a general model which assumes that the cell residence times have general non-lattice distribution, and the call holding times are distributed with general distributions such as Gamma, hyper-exponential and hyper-Erlang distributions. While the advantages of such distributions are clear in that they reflect the emerging services, applying these distributions to the model is a non-trivial task.

The following general assumptions will be used throughout this research:

- The mobile users are uniformly scattered in the whole service area of the PCS network, i.e., the PCS network is homogeneous;
- The new call arrivals forms a Poisson process;
- The cell residence times are independent, identically distributed (iid) with non-lattice distribution;

- The call holding times are independent, identically distributed (iid) with non-lattice distribution.

These assumptions are general enough to capture some of the essence of PCS network traffic while enable us to obtain analytical results for the performance of PCS networks. As a final note, we discuss the residual life (r_1) of call holding time in the originating cell. Throughout the dissertation, we use the assumption that the Residual Life Theorem is valid for our PCS models. It can observed that when the residual life has general distribution, all results can be easily modified.

1.6 Organization of This Dissertation

This dissertation is organized as follows. In the next chapter, we study the cell traffic and handoff rate, we present necessary and sufficient conditions for two commonly used assumptions in the literature to be valid, which clarify under what conditions the Erlang-B formula or the open queueing network theorem for computing blocking probabilities can be used. We also give analytical expressions for both channel occupancy times and handoff rate. In Chapter 3, we demonstrate how to analytically compute the call completion probability and the expected effective call holding times under fairly general assumptions on the call holding times (exponentially or Erlang distributed) and cell residence times. Chapter 4 continues the research on Chapter 3; we generalize the analytical results to deal with the cases where the call holding times are distributed with more general distributions. The results provide more models of PCS networks to choose for the field trials. Moreover, billing rating plans based on the effective call holding times are proposed. Chapter 5 focus on a few billing strategies and performance analyses for PCS networks. Our strategies will facilitate the billing experts to monitor and upgrade the billing systems and to detect possible fraudulent activities. In the last chapter, we conclude this dissertation with future research direction.

Chapter 2

CELL TRAFFIC, CHANNEL OCCUPANCY TIMES AND HANDOFF RATE

2.1 Introduction

A convergence of Personal Communications Services (PCS) has led to the emergence of networks poised to provide integrated services such as voice and data to mobile users anywhere, anytime ([34, 6]) in an uninterrupted and seamless way, using advanced microcellular and handoff concepts ([29]). Due to the growing importance of PCS it is necessary to study their behavior for performance evaluation, optimization, management and billing, under realistic conditions. Two of the key entities in evaluating these systems are the channel occupancy distribution and the handoff rates.

The channel occupancy time distribution has been studied in the past quite extensively for classical cellular systems. A common assumption in these studies has been that the call holding times (the times requested for a call connection) are exponentially distributed. Hong and Rappaport ([24]) proposed a traffic model for cellular mobile radio telephone systems in which they analyzed the performance and the channel occupancy time distribution approximated by exponential distribution under the assumption of exponential call holding times. Using a simulation model, Guerin ([20]) showed that for some cases the channel occupancy time distribution is quite “close” to exponential distribution, and that for “low” rate of change of direction the channel occupancy time distribution displays rather poor agreement with the exponential distribution. Lin et al ([40]) showed that when the call holding times and cell residence times are exponentially distributed, then the channel occupancy time is also exponentially distributed. In ([36]) Lin et al obtained the expected channel occupancy times for both new calls and handoff calls. Nanda ([50]) studied the handoff rate under the assumption that the call holding times are exponentially distributed and there are no blocking and forced terminations (i.e., corresponding to the ideal assumption that there is an infinite number of channels available in each cell).

To model the PCS networks in a realistic way several observations are in order. First, due to the wide spectrum of the integrated communications services (such as phone calls, information retrievals, Internet surfing, etc.) carried jointly over a PCS network, the call holding times may not be exponentially distributed, and assumption made in the past for evaluating the behavior of classical wireless or wireline telephone networks. Secondly, due to user mobility (portables or mobile computers) and the irregular geographical cell structures, the cell residence times (the time a user spends in a cell) will also typically have a general distribution. Thirdly, the channel occupancy time in a cell, i.e. the time the channel is occupied by a call in a cell (a new call, a handoff call, and regardless of the call being completed in the cell or moving out of the cell) is also not necessarily distributed exponentially as generally assumed in the past ([40, 50, 71, 72]). Lastly, to complete a call connection in a PCS network, the user may traverse several cells and the call may consequently be handed off many times before it is completed.

This chapter presents a systematic study of the channel occupancy times and handoff rate in PCS systems under real systems assumptions, leading to a number of new results. Under the assumption that the call holding time is exponentially distributed, we present a necessary and sufficient condition for the new call channel occupancy time (the handoff call channel occupancy time, or the channel occupancy time) to be exponentially distributed. Specifically, we show that the new call channel occupancy time is exponentially distributed if and only if cell residence times are exponentially distributed. When cell residence times are not exponentially distributed, we derive formulae to compute the distribution of channel occupancy times. In order to apply Erlang-B formula, the cell arrival traffic has been assumed to be Poisson, a common assumption for the computation of blocking probability in cellular systems ([8, 18, 40, 66, 71, 72]). However, for the cell arrival traffic which consists of merged traffic of new calls and handoff calls, we show that the merged traffic is Poisson if and only if the cell residence times are exponentially distributed.

The average number of handoffs during a call connection (also referred to as the handoff rate) is the second focus of this chapter. Since service quality and cost are tied to the handoff rate and the handoff call arrival rate, the expected number of handoffs (i.e. the handoff rate) becomes an important performance index for the PCS networks and consequently needs to be known. In this chapter we, therefore, derive a general formula for the computation of the handoff rate under general conditions, i.e., when the call holding times and cell residence times are generally distributed with non-lattice distribution functions. Using the handoff rate, it becomes possible to then compute the handoff call arrival rate.

Lastly, we observe that the technique proposed in this chapter also leads to a practical model for deriving performance evaluation of GSM based mobile computing systems. GSM provides data capabilities to support mobile computing applications. The data capabilities include GSM phase 2 bearer data service, phase 2+ high speed circuit switched data (HSCSD), and general packet radio service (GPRS) ([10, 45]).

These data services are either circuit switched or packet switched at the transport layer (i.e., at the wireline backbone network), but are all circuit switched at the radio interface. In other words, a mobile computing user will occupy a traffic channel just like a typical voice user. Thus, the modeling of radio resource allocation for GSM-supported mobile computing system is similar to the radio channel allocation for PCS, except that the computing session times in mobile computing are unlikely to be exponentially distributed as in traditional PCS systems. As a result, previously proposed approaches ([8, 42, 40, 66, 71, 72]) are not appropriate for modeling the mobile computing systems (due to the exponential call holding time assumption). By allowing the use of general call holding time distribution, our model can be used to study GSM based mobile computing systems with arbitrary computing session periods.

This chapter is organized as follows. In the next section, we discuss the properties of channel occupancy times and the merged cell traffic. In the third section we present analytical formulae for the computation of channel occupancy time distributions for general cell residence times. In the fourth section, we present results on handoff rate calculation. An illustrative example is given in the fifth section, with the last section concluding the chapter.

2.2 Comments on Classical Assumptions Regarding Handoff Traffic and Channel Occupancy Times

In order to find the blocking probability in PCS networks, a commonly used assumption ([8, 18, 66, 72]) is that the channel occupancy times are exponentially distributed. While this is a reasonable assumption for wired telephone traffic and most currently used cellular networks, we show that this assumption holds only for the case

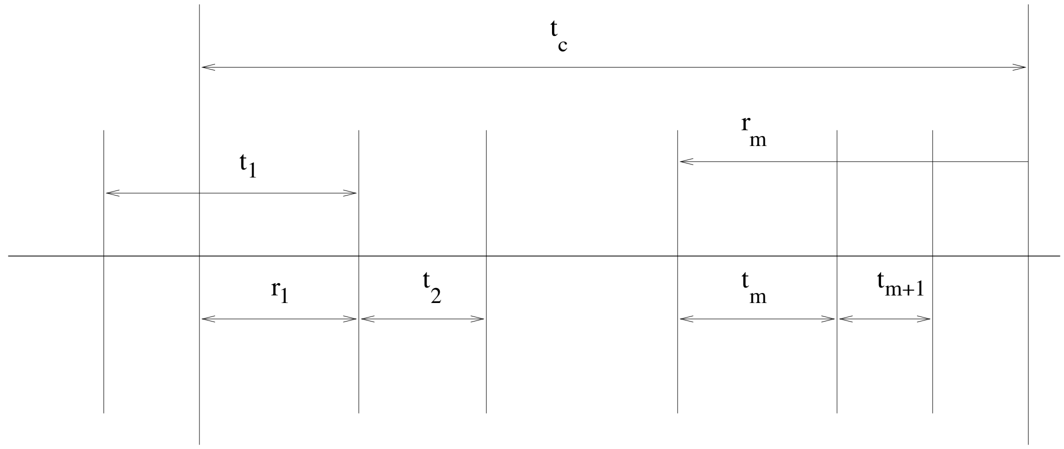


Figure 2.1: The time diagram for call holding time and cell residence time

when cell residence times are exponentially distributed, a property which does not hold for emerging PCS networks. The second commonly used assumption is that the merged traffic from the new call traffic and handoff traffic in a cell is a Poisson process ([8, 18, 40, 66, 71, 72]). This assumption is needed to apply the well-known Erlang-B formula (or its extended version) using the $M/M/c/c$ (or $M/G/c/c$) queueing model to compute the blocking probability in a cell. Although this assumption regarding merged traffic properties may be a good approximation for some cases, it can not be expected to be appropriate for most PCS applications. In fact the fit of the exponential distribution for these applications has not been quantified in the literature until now. In here we show that this assumption is valid only for the case when cell residence times are exponentially distributed.

Let us consider the first assumption first. Assume that the call holding times (the times of requested connections to a PCS network for new calls) are exponentially distributed with parameter μ. Let t_c be the call holding time for a typical new call, t_m be the cell residence time, r_1 be the time between the instant the new call is initiated at and the instant the new call moves out the cell if the new call is not completed, r_m $(m > 1)$ be the residual life of call holding time when the call finishes $m-$th handoff successfully. Let t_{no} and t_{ho} denote the channel occupancy times for a new call and a handoff call, respectively. Then, from Figure 2.1, for a new call, the channel occupancy time will be

$$t_{no} = \min\{t_c, r_1\}, \tag{2.1}$$

and the channel occupancy time is

$$t_{ho} = \min\{r_m, t_m\}. \tag{2.2}$$

Let t_c, t_m, r_1, t_{ho} and t_{no} have density functions $f_c(t)$, $f(t)$, $f_r(t)$, $f_{ho}(t)$ and $f_{no}(t)$ with their corresponding Laplace transforms $f_c^*(s)$, $f^*(s)$, $f_r^*(s)$, $f_{ho}^*(s)$ and $f_{no}^*(s)$, respectively.

We next show that the handoff call channel occupancy time t_{ho} is exponentially distributed if and only if the cell residence time t_m is exponentially distributed. From (2.2), we obtain the probability

$$\begin{aligned}
\Pr(t_{ho} \le t) &= \Pr(r_m \le t \text{ or } t_m \le t) = \Pr(r_m \le t) + \Pr(t_m \le t) - \Pr(r_m \le t, t_m \le t) \\
&= \Pr(r_m \le t) + \Pr(t_m \le t) - \Pr(r_m \le t)\Pr(t_m \le t) \\
&= \Pr(t_c \le t) + \Pr(t_m \le t) - \Pr(t_c \le t)\Pr(t_m \le t), \qquad (2.3)
\end{aligned}$$

where we have used the independency of r_m and t_m, and memoryless property of the exponential distribution of t_c from the distribution of r_m has the same distribution as t_c (either from the Residual Life Theorem or the argument in [36]). Differentiating (2.3), we obtain

$$\begin{aligned}
f_{ho}(t) &= f_c(t) + f(t) - f_c(t)\Pr(t_m \le t) - \Pr(t_c \le t) f(t) \\
&= f_c(t)\int_t^\infty f(\tau)d\tau + f(t)\int_t^\infty f_c(\tau)d\tau. \qquad (2.4)
\end{aligned}$$

Suppose that the cell residence times are exponentially distributed with parameter η, then from (2.4) we obtain

$$f_{ho}(t) = \mu e^{-\mu t} \times e^{-\eta t} + \eta e^{-\eta t} \times e^{-\mu t} = (\mu + \eta)e^{-(\mu+\eta)t},$$

which is an exponential distribution. Conversely, suppose that the handoff channel occupancy time has exponential distribution with parameter γ, let $Y(t) = \int_t^\infty f(t)$, then $\dot{Y}(t) = -f(t)$ (overdot $\cdot$ denotes the differentiation symbol). From (2.4), we obtain

$$\mu e^{-\mu t} Y(t) + e^{-\mu t} f(t) = \gamma e^{-\gamma t},$$

i.e.,

$$\dot{Y}(t) = \mu Y(t) - \gamma e^{-(\gamma-\mu)t},$$

from which we obtain

$$\begin{aligned} Y(t) &= e^{\mu t} Y(0) + \int_0^t e^{\mu(t-\tau)} [-\gamma e^{-(\gamma-\mu)\tau}] d\tau \\ &= e^{\mu t} \left\{ Y(0) - \gamma \int_0^t e^{-\gamma \tau} d\tau \right\} = e^{-(\gamma-\mu)t}. \end{aligned}$$

Thus, $f(t) = -\dot{Y}(t) = (\gamma - \mu) e^{-(\gamma-\mu)t}$, i.e., the cell residence time must be exponentially distributed.

We now consider the new call channel occupancy time distribution case. From (2.1) and a similar argument, we obtain

$$f_{no}(t) = f_c(t) \int_t^\infty f_r(\tau) d\tau + f_r(t) \int_t^\infty f_c(\tau) d\tau.$$

Suppose that the new call cell channel occupancy time is exponentially distributed with parameter, say, μ_1, from this identity and a similar argument as for the handoff call channel occupancy time case, we can deduce that $f_r(t) = (\mu_1 - \mu) e^{-(\mu_1-\mu)t}$, which is also an exponential distribution. Let $F(t)$ denote the distribution function of the cell residence time with mean $1/\eta$, from the Residual Life Theorem ([32, 46]) we have

$$f_r(t) = \eta(1 - F(t)),$$

from which we obtain

$$F(t) = 1 - \frac{\mu_1 - \mu}{\eta} e^{-(\mu_1-\mu)t}.$$

From this and $F(0) = 0$, we obtain $\mu_1 - \mu = \eta$, so $F(t) = 1 - e^{-\eta t}$, and we conclude that the cell residence times are also exponentially distributed. This shows that for an

PCS network with exponential call holding times, the new call channel occupancy time is exponentially distributed if and only if the cell residence times are exponentially distributed.

In the preceding discussion we separated calls into new calls and handoff calls, when considering the channel occupancy times. If such distinction is not made, then we need to consider the channel occupancy time distribution for any call (either new call or handoff call), i.e., the channel occupancy time for the merged traffic of new calls and handoff calls, as used in current literature. We will simply call this the channel occupancy time without any modifiers such as new call or handoff call. Let t_{co} denote the channel occupancy time and λ_h the handoff call arrival rate (which will be discussed in a later section). Then, it is easy to show that $t_{co} = t_{no}$ with probability $\lambda/(\lambda+\lambda_h)\stackrel{\Delta}{=}p$ and $t_{co} = t_{ho}$ with probability $\lambda_h/(\lambda+\lambda_h)\stackrel{\Delta}{=}q$. Let $f_{co}(t)$ and $f^*_{co}(s)$ be its density function and the corresponding Laplace transform. It is easy to obtain

$$\begin{aligned} & f_{co}(t) = pf_{no}(t) + qf_{ho}(t) \\ = \; & pf_c(t)\int_t^\infty f_r(\tau)d\tau + pf_r(t)\int_t^\infty f_c(\tau)d\tau + qf_c(t)\int_t^\infty f(\tau)d\tau + qf(t)\int_t^\infty f_c(\tau)d\tau \\ = \; & e^{-\mu t}\left\{p\mu\eta\int_t^\infty[1-F(\tau)]d\tau + (p\eta+q\mu)[1-F(t)] + qf(t)\right\}. \end{aligned} \tag{2.5}$$

It is straightforward that when the cell residence times are exponentially distributed, then the channel occupancy time is also exponentially distributed. Conversely, suppose that the channel occupancy time is exponentially distributed with, say, parameter γ, i.e., $f_{co}(t) = \gamma e^{-\gamma t}$, from (2.5) we obtain

$$p\mu\eta\int_t^\infty[1-F(\tau)]d\tau + (p\eta+q\mu)[1-F(t)] + qf(t) = \gamma e^{-(\gamma-\mu)t}. \tag{2.6}$$

From the left hand side of equation (2.6) with the properties of the distribution function, we can deduce that $\gamma-\mu \geq 0$. Let $Y(t) = \int_t^\infty[1-F(\tau)]d\tau$, then $F(t) = 1+\dot{Y}(t)$, $\ddot{Y}(t) = f(t)$, taking these into (2.6), we obtain

$$\ddot{Y}(t) - (\frac{p}{q}\eta+\mu)\dot{Y}(t) + \frac{p}{q}\eta\mu Y(t) = \frac{\gamma}{q}e^{-(\gamma-\mu)t}. \tag{2.7}$$

One particular solution of (2.7) is in the form

$$Y_1(t) = Be^{-(\gamma-\mu)t},$$

where

$$B = \frac{\gamma}{q\gamma^2 + \gamma(p\eta + q\mu) + p\mu\eta} > 0.$$

Noticing that the characteristic equation of (2.7) is

$$s^2 - (\frac{p}{q}\eta + \mu)s + \frac{p}{q}\eta\mu = (s-\mu)(s - \frac{p}{q}\eta) = 0,$$

which has roots $s = \mu > 0$ and $s = (p/q)\eta > 0$. If these two roots are equal, then all solutions of (2.7) are given by

$$Y(t) = C_1 e^{\mu t} + C_2 t e^{\mu t} + Be^{-(\gamma-\mu)t},$$

where C_1 and C_2 are constants. Since $\lim_{t\to\infty} Y(t) = 0$, we must have $C1 = C_2 = 0$, so $Y(t) = Be^{-(\gamma-\mu)t}$. Similarly, if the two roots are not equal, then all solutions of (2.7) are

$$Y(t) = C_1 e^{\mu t} + C_2 e^{-[(p/q)\eta]t} + Be^{-(\gamma-\mu)t},$$

so $C_1 = C_2 = 0$ from the definition of $Y(t)$. In any case, $Y(t)$ must be in the form $Y(t) = Be^{-(\gamma-\mu)t}$. So, we have

$$F(t) = 1 + \dot{Y}(t) = 1 - B(\gamma-\mu)e^{-(\gamma-\mu)t}.$$

From $F(0) = 0$, we obtain $B(\gamma-\mu) = 1$, hence $F(t) = 1 - e^{-(\gamma-\mu)t}$, which implies that the cell residence times are exponentially distributed. In summary, we have shown that the channel occupancy time is exponentially distributed if and only if the cell residence times are exponentially distributed.

Next, we discuss the second commonly used assumption, i.e., the merged traffic of new calls and handoff calls in a cell is Poissonian. Assume that in a typical cell of a PCS network, the new call arrivals are Poisson, then the handoff arrivals to the cell are independent of the new call arrivals. Let $N_n(t)$ and $N_h(t)$ be the numbers of new calls and handoff calls, respectively, up to time t. Let $N(t)$ be the number of calls from the merged traffic of the new call traffic and the handoff traffic. Then, we have

$$N(t) = N_n(t) + N_h(t), \tag{2.8}$$

We want to use the $Z-$transform theory and the following result ([32]): for a traffic with counting process $N(t)$, $N(t)$ is a Poisson process if and only if its $Z-$transform $E[z^{N(t)}]$ is equal to $\exp(-\lambda t(1-z))$. If $N_h(t)$ is a Poisson process, then obvious $N(t)$ is a Poisson process, i.e., the merged traffic is a Poisson process. Suppose that $N(t)$ is a Poisson process with parameter λ_m and $N_n(t)$ is a Poisson process with parameter λ, then from (2.8) we have

$$E[z^{N(t)}] = e^{-\lambda_m t(1-z)} = E[z^{N_n(t)+N_h(t)}] = E[z^{N_n(t)}]E[z^{N_h(t)}] = e^{-\lambda t(1-z)}E[z^{N_h(t)}].$$

From this, we obtain

$$E[z^{N_h(t)}] = e^{-(\lambda_m-\lambda)t(1-z)}.$$

Thus, the handoff traffic $N_h(t)$ is also a Poisson process. It is well-known ([5, 32, 46]) that for the $M/G/c$ queueing system with first-come-first-serve (FCFS) strategy, the departure process is a Poisson process if and only if the service time is exponentially distributed. We next observe that the handoff traffic is the departure process of the queueing system with the Poisson arrivals (the merged traffic) and with two virtual "servers": one "server" for the calls which complete the connection successfully in the cell (the service time distribution is exponentially distributed due to the memoryless property of exponential distribution), the other "server" for calls which need handoffs (which forms the handoff traffic). The departure process from the first server is a Poisson process since the departure process from the $M/M/1$ queue is a Poisson process ([5]). As the handoff traffic is Poisson from the above discussion, from the earlier referenced Burke's result, the channel occupancy time must be exponentially distributed, hence the cell residence times must be exponentially distributed too.

Summarizing the preceding discussion, we finally obtain:

Theorem 2.1. For a PCS network with exponential call holding times and Poisson new call arrivals, we can state:

(1) the new call channel occupancy time is exponentially distributed if and only if the cell residence times are exponentially distributed,

(2) the handoff call channel occupancy time is exponentially distributed if and only if the cell residence times are exponentially distributed,

(3) the channel occupancy time is exponentially distributed if and only if the cell residence times are exponentially distributed, and

(4) the merged traffic from the new call traffic and handoff call traffic is still Poissonian if and only if the cell residence times are exponentially distributed.

2.3 Channel Occupancy Times

In the preceding section we concluded that for channel occupancy time to be exponentially distributed, the cell residence times had to be exponentially distributed. However, the assumption of exponential cell residence times is too restrictive, since it is important to know the distribution of channel occupancy time for generally distributed cell residence times. It is furthermore important to find out how "close" is the exponential distribution to the distribution of channel occupancy times. We address these issues next.

From (2.4), applying Laplace transform, we obtain

$$
\begin{aligned}
f_{ho}^*(s) &= f^*(s) + f_c^*(s) - \int_0^\infty e^{-st} f(t) \int_0^t f_c(\tau) d\tau dt - \int_0^\infty e^{-st} f_c(t) \int_0^t f(\tau) d\tau dt \\
&= f^*(s) + f_c^*(s) - \mu \int_0^\infty e^{-(s+\mu)t} \int_0^t f(\tau) d\tau dt - \int_0^\infty e^{-st}(1 - e^{-\mu t}) f(t) dt \\
&= \frac{\mu}{s+\mu} + \frac{s}{s+\mu} f^*(s+\mu).
\end{aligned}
$$

From this, it is easy to obtain that the expected handoff channel occupancy time is given by (we will use $h^{(i)}(x)$ to denote the ith derivatives of any function $h(x)$ at point x in the subsequent development)

$$E[t_{ho}] = -f_{ho}^{*(1)}(0) = \frac{1}{\mu}(1 - f^*(\mu)).$$

Similarly, from (2.1), we obtain

$$f_{no}^*(s) = \frac{\mu}{s+\mu} + \frac{s}{s+\mu} f_r^*(s+\mu). \tag{2.9}$$

Since r_1 is the residual life of the cell residence time, from the Residual Life Theorem ([32, 46]) we have

$$f_r^*(s) = \frac{\eta}{s}[1 - f^*(s)],$$

where $\eta = 1/E[t_m]$, i.e., the cell residence rate. Taking this into (2.9), we obtain

$$f_{no}^*(s) = \frac{\mu}{s+\mu} + \frac{\eta s}{(s+\mu)^2}[1 - f^*(s+\mu)],$$

from which we also obtain the expected new call channel occupancy time

$$E[t_{no}] = -f_{no}^{*(1)}(0) = \frac{1}{\mu} - \frac{\eta}{\mu^2}[1 - f^*(\mu)].$$

Similarly, we can obtain formulae for channel occupancy time.

It is commonly assumed that the new call channel occupancy time and the handoff call channel occupancy time have the same distribution. However, we claim that this is true only when the cell residence times are exponentially distributed. In fact, suppose that channel occupancy times for both new call and handoff call have the same distribution, then their Laplace transforms are equal, $f_{no}^*(s) = f_{ho}^*(s)$, we obtain that $f^*(s+\mu) = \eta/(s+\mu+\eta)$, hence $f^*(s) = \eta/(s+\eta)$, this implies that the cell residence times must be exponentially distributed.

Summarizing the above discussions, we arrive at:

Theorem 2.2. For a PCS network with exponential call holding times and Poisson new call arrivals with arrival rate λ, we have:

(1) the Laplace transform of the density function of the new call channel occupancy time is given by

$$f^*_{no}(s) = \frac{\mu}{s+\mu} + \frac{\eta s}{(s+\mu)^2}[1 - f^*(s+\mu)], \tag{2.10}$$

and the expected new call channel occupancy time is

$$E[t_{no}] = \frac{1}{\mu} - \frac{\eta}{\mu^2}[1 - f^*(\mu)]; \tag{2.11}$$

(2) the Laplace transform of the density function of the handoff call channel occupancy time is given by

$$f^*_{ho}(s) = \frac{\mu}{s+\mu} + \frac{s}{s+\mu} f^*(s+\mu), \tag{2.12}$$

and the expected handoff call channel occupancy time is

$$E[t_{ho}] = \frac{1}{\mu}(1 - f^*(\mu)). \tag{2.13}$$

(3) let λ_h denote the handoff call arrival rate, then the Laplace transform of the density function of channel occupancy time is given by

$$f^*_{co}(s) = \frac{\lambda}{\lambda+\lambda_h} f^*_{no}(s) + \frac{\lambda_h}{\lambda+\lambda_h} f^*_{ho}(s). \tag{2.14}$$

and the expected channel occupancy time is given by

$$E[t_{co}] = \frac{1}{\mu} - \frac{\lambda\eta}{(\lambda+\lambda_h)\mu^2}\left[1 - \left(1 - \frac{\lambda_h\mu}{\lambda\eta}\right) f^*(\mu)\right]. \tag{2.15}$$

(4) the new call channel occupancy time and the handoff call occupancy time have the same distribution if and only if the cell residence times are exponentially distributed.

Remark: The expected channel occupancy times $E[t_{no}]$ and $E[t_{ho}]$ given in (2.11) and (2.13) have also been obtained in [36] from a direct approach.

Using Theorem 2.2, we can obtain the channel occupancy time density functions by the inverse Laplace transform, from which the distribution functions can be obtained. In order to observe how "close" the exponential approximations can be, we have to determine which exponential distribution functions should be chosen. It is known that an exponential distribution is uniquely determined by its expected value. It is reasonable to use the exponential distribution whose expected value is equal to the real expected value of channel occupancy time (either (2.11), (2.13) or (2.15)). There are many criteria to evaluate this approximation, known in statistics as the "goodness of fit". A good choice will be the distance between the distribution functions of the real data and the exponential data. Hong and Rappaport ([24]) proposed the following measure for the "goodness of fit"

$$G = \frac{\int_0^\infty |F(t) - (1 - e^{-at})| \, dt}{2 \int_0^\infty (1 - F(t)) dt},$$

where $F(t)$ is the distribution function of real data and a is the expected value of the exponential distribution used for the approximation. This measure is the normalized accumulated difference of distribution functions. Comparisons can also be done graphically by drawing distribution functions, which will be used in the illustrative example in fifth section.

2.4 Handoff Rate for General Call Holding Times and Cell Residence Times

As pointed out in the Introduction it is necessary to study the number of handoffs occurring in a call if one wishes to evaluate a PCS network. For certain special cases, results for handoff rate exist in the current literature. In [50], Nanda presented an analytic result of handoff rate for the case when the call holding time is exponentially distributed and no handoff failure occurs (which is equivalent to the case where each cell has an infinite number of channels available). This is, of course, the ideal case. Lin et al ([40]) considered the more practical case where handoff failures are taken into consideration and presented a formula for the case when call holding times are

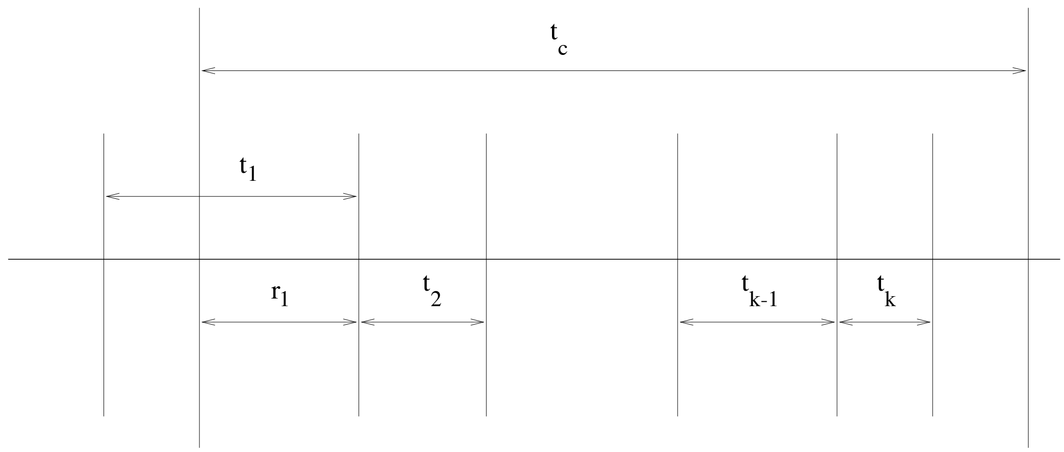


Figure 2.2: The time diagram for k handoffs

distributed exponentially. In this section, we present a formula for the general case where the call holding times and cell residence times are generally distributed, and handoff failures are accounted for.

Consider a typical call. Let $t_1, t_2, \cdots$ denote the cell residence times and r_1 denote the residual life of the new call (i.e., the time interval between the call arrival and the exit of the cell of the portable). Let t_c denote the call holding time for the typical call. We use Figure 2.2 to indicate the time diagram for k handoffs. Let H be the number of handoffs of a typical nonblocking call (either completed or forced to terminate) during the call connection. We can now study the property of the number H of handoffs under a general call holding time and cell residence time distributions.

Let $f^*(s)$ and $f_c^*(s)$ be defined as in the last section. We will assume that $t_1, t_2, \ldots$ have independently and identically distributed (iid) nonlattice distributions, and that the call holding time has generally distributed nonlattice distribution ([46]). The following is apparent: $H = 0$ if and only if the call is not blocked and the call holding time t_c is shorter than the residual life r_1, i.e., the call completes before the portable moves out of the cell; $H = 1$ if and only if the call is not blocked initially, then it either makes a successful handoff and completes the call successfully in the new cell, or is forced to terminate because of the first handoff failure, and so on. If the blocking probability for a new call is p_o and the probability for a handoff call to be forced to terminate is p_f, then we have

$$\begin{aligned}
\Pr(H=0) &= (1-p_o)\Pr(r_1 \geq t_c) \\
\Pr(H=1) &= (1-p_o)\Pr(r_1 < t_c \leq r_1+t_2)(1-p_f) + (1-p_o)\Pr(t_c > r_1+t_2)p_f \\
&\vdots \\
\Pr(H=k) &= (1-p_o)\Pr(r_1+t_2+\cdots+t_k < t_c \leq r_1+t_2+\cdots+t_{k+1})(1-p_f)^k \\
&\quad + (1-p_o)\Pr(t_c > r_1+t_2+\cdots+t_k)(1-p_f)^{k-1}p_f \\
&\vdots
\end{aligned} \tag{2.16}$$

Let σ be a positive number which will be appropriately chosen for Laplace inverse transform.

We first calculate $\Pr(H=0)$. Since the Laplace transform of $\int_t^\infty f_r(\tau)d\tau$ is $(1-f_r^*(s))/s$ and the independency of t_c and r_1, from the first equation in (2.16) and the inverse Laplace transform we have

$$\begin{aligned}
\Pr(H=0) &= (1-p_o)\int_0^\infty \Pr(r_1 \geq t) f_c(t)dt \\
&= (1-p_o)\int_0^\infty \int_t^\infty f_r(\tau)d\tau f_c(t)dt \\
&= (1-p_o)\int_0^\infty \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f_r^*(s)}{s} e^{st} ds f_c(t)dt \\
&= \frac{1-p_o}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f_r^*(s)}{s}\int_0^\infty f_c(t)e^{st}dtds \\
&= \frac{1-p_o}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f_r^*(s)}{s} f_c^*(-s)ds.
\end{aligned} \tag{2.17}$$

Next, we compute $\Pr(H=k)$ for $k>0$. Before we do that, we need to compute $Pr(r_1+t_2+\cdots+t_k \leq t_c)$. Let $\xi = r_1+t_2+\cdots+t_k$. Let $f_\xi(t)$ and $f_\xi^*(s)$ be the density function and the Laplace transform of ξ. From the independency of $r_1, t_2, t_3, \cdots$, we have

$$f_\xi^*(s) = E[e^{-s\xi}] = E[e^{-sr_1}]\prod_{i=2}^{k} E[e^{-st_i}] = f_r^*(s)(f^*(s))^{k-1}.$$

So the density function is given by

$$f_\xi(t) = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} f_r^*(s)(f^*(s))^{k-1}e^{st}ds.$$

Also, the Laplace transform of $\Pr(\xi \le t)$ (the distribution function) is $f_\xi^*(s)/s$. Thus, we have

$$\begin{aligned}
\Pr(r_1+t_2+\cdots+t_k \le t_c) &= \int_0^\infty \Pr(\xi \le t) f_c(t)dt \\
&= \int_0^\infty \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[f^*(s)]^{k-1}}{s} e^{st} ds f_c(t)dt \\
&= \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[f^*(s)]^{k-1}}{s} f_c^*(-s)ds.
\end{aligned}$$

Taking this into (2.16), we obtain

$$\begin{aligned}
&\Pr(H=k) = (1-p_o)[\Pr(t_c \ge r_1+t_2+\cdots+t_k) \\
&\quad - \Pr(t_c \ge r_1+t_2+\cdots+t_{k+1})](1-p_f)^k \\
&\quad + (1-p_o)\Pr(t_c \ge r_1+t_2+\cdots+t_k)(1-p_f)^{k-1}p_f \\
&= \frac{(1-p_o)(1-p_f)^k}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[f^*(s)]^{k-1}[1-f^*(s)]}{s} f_c^*(-s)ds \\
&\quad + \frac{(1-p_o)(1-p_f)^{k-1}p_f}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[f^*(s)]^{k-1}}{s} f_c^*(s)ds \\
&= \frac{(1-p_o)(1-p_f)^{k-1}}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[1-(1-p_f)f^*(s)][f^*(s)]^{k-1}}{s} f_c^*(-s)ds.
\end{aligned} \tag{2.18}$$

Now we want to find the $Z-$transform (moment generating function) for the number of handoffs. Let $H(z)$ be the moment generating function, from (2.17) and (2.18) we have

$$\begin{aligned}
&H(z) = E[z^H] = \sum_{k=0}^\infty z^k \Pr(H=k) \\
&= \Pr(H=0) + \sum_{k=1}^\infty z^k \Pr(H=k) \\
&= \Pr(H=0) + \\
&\quad + \frac{(1-p_o)z}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{f_r^*(s)[1-(1-p_f)f^*(s)]}{s}\left(\sum_{k=1}^\infty [(1-p_f)zf^*(s)]^{k-1}\right) f_c^*(-s)ds
\end{aligned}$$

$$
\begin{aligned}
&= \frac{(1-p_o)}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f_r^*(s)}{s}f_c^*(-s)ds \\
&\quad +\frac{(1-p_o)z}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{f_r^*(s)[1-(1-p_f)f^*(s)]}{s[1-(1-p_f)zf^*(s)]}f_c^*(-s)ds \\
&= \frac{1-p_o}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f_r^*(s)+z[f_r^*(s)-(1-p_f)f^*(s)]}{s[1-(1-p_f)zf^*(s)]}f_c^*(-s)ds \\
&= \frac{1-p_o}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{s-\eta[1-f^*(s)]+z[\eta(1-f^*(s))-(1-p_f)sf^*(s)]}{s^2[1-z(1-p_f)f^*(s)]}f_c^*(-s)ds,
\end{aligned}
\tag{2.19}
$$

where $f_r^*(s)=\eta(1-f^*(s))/s$. It is obvious that when $|z|\le 1$, the integrand without term $f_c^*(-s)$ is analytic on the right half open complex plane. If $f_c^*(-s)$ has no branch point and has only finite possible isolated singular points in the right half plane (which is equivalent to saying that $f_c^*(s)$ has only finite possible isolated singular point in the left half plane), then the Residue Theorem can be applied to (2.19) using a semicircular contour in the right half plane. Indeed, if we use σ_c denote the singular points of $f_c^*(-s)$ in the right half complex plane, then from (2.19) and the Residue Theorem ([35]) we obtain

Theorem 2.3. If the density function of the iid calling holding times has only finite possible isolated singular points, then the moment generating function for the number of handoffs is given by

$$
H(z)=-(1-p_o)\sum_{p\in\sigma_c}\operatorname*{Res}_{s=p}\frac{s-\eta[1-f^*(s)]+z[\eta(1-f^*(s))-(1-p_f)sf^*(s)]}{s^2[1-z(1-p_f)f^*(s)]}f_c^*(-s),
\tag{2.20}
$$

where $\operatorname*{Res}_{p}$ denotes the residue at singular point $s=p$.

Proof: Choosing such σ that all singular points of $f_c^*(-s)$ in the right half plane are on the right of vertical line $s=\sigma$, and choosing the contour enclosed by the semicircle at center $s=\sigma+j0$ and with radius sufficiently large, then we can apply the Residue Theorem to complete the proof.

We can apply Theorem 2.3 to obtain the moments of the number of handoffs. However, we can apply (2.19) to easily find the expected number of handoffs. Differentiating $H(z)$ at $z = 1$, we obtain

$$E[H] = \frac{1-p_o}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{\eta[1 - f^*(s)]}{s^2[1 - (1-p_f)f^*(s)]} f_c^*(-s) ds. \tag{2.21}$$

Applying the Residue Theorem, we obtain

Theorem 2.4. If the density function of the iid calling holding times has only finite possible isolated singular points, then the expected number of handoffs for a nonblocking call (handoff rate) is given by

$$E[H] = -\eta(1-p_o) \sum_{p \in \sigma_c} \operatorname*{Res}_{s=p} \frac{1 - f^*(s)}{s^2[1 - (1-p_f)f^*(s)]} f_c^*(-s). \tag{2.22}$$

If the call holding times are exponentially distributed with parameter μ, then $f_c^*(-s) = \mu/(-s+\mu)$, which has a unique singular point, and $\sigma_c = \{\mu\}$. From Theorem 2.4, we obtain

$$E[H] = (1-p_o) \frac{\eta(1 - f^*(\mu))}{\mu[1 - (1-p_f)f^*(\mu)]}. \tag{2.23}$$

This has been obtained in [36] and [40] using different approaches.

If there is no blocking and no forced termination (the ideal case when there are infinitely many number of channels available in each cell), then $p_o = p_f = 0$, and from (2.23) $E[H] = \eta/\mu$, or the handoff rate is $1/E[H] = \mu/\eta$, which is obtained in [50]. If there are handoff failures, the expected number of handoffs is intuitively smaller. Indeed, for any handoff schemes, since $1 - f^*(\mu) \leq 1 - (1-p_f)f^*(\mu)$ for any p_f $(0 \leq p_f \leq 1)$, we have from (2.23)

$$E[H] \leq (1-p_o)\frac{\eta}{\mu} \leq \frac{\eta}{\mu}.$$

Let us assume that the call holding times are iid with Erlang distribution

$$f_c(t) = \frac{\alpha^m t^{m-1}}{(m-1)!} e^{-\alpha t}, \;\; f_c^*(s) = \left(\frac{\alpha}{s+\alpha}\right)^m, \tag{2.24}$$

where $\alpha = m\eta$ is the scale parameter and m is a positive integer. When $m = 1$, it gives the exponential distribution. We can also find a simple formula for the handoff rate. Let

$$g(s) = \frac{1 - f^*(s)}{s^2[1 - (1 - p_f)f^*(s)]}. \tag{2.25}$$

Then, $f_c^*(-s) = [\alpha/(-s+\alpha)]^m$ has a unique singular point $\sigma_c = \{\alpha\}$, from Theorem 2.4, we obtain

$$E[H] = \eta(-1)^{m+1}\alpha^m(1 - p_o)\frac{g^{(m-1)}(\alpha)}{(m-1)!}, \quad \alpha = m\mu. \tag{2.26}$$

For this case, if there is no handoff failure, i.e., $p_o = p_f = 0$, from (2.25), we have $g(s) = 1/s^2$ and $g^{(m-1)}(\alpha) = (-1)^{m-1}m!\alpha^{-(m+1)}$. From (2.26), we have

$$E[H] = \frac{m\eta}{\alpha} = \frac{\eta}{\mu}.$$

In fact, it can be shown that for the ideal case ($p_o = p_f = 0$), the expected number of handoffs for any cell residence time distribution and any call holding time distribution is $E[H] = \eta/\mu$. Indeed, from (2.21), if $p_f = 0$, we obtain

$$\begin{aligned} E[H] &= \frac{(1-p_o)\eta}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1}{s^2}f_c^*(-s)ds = \eta(1-p_o)\int_0^\infty f_c(t)\left[\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1}{s^2}e^{st}ds\right]dt \\ &= \eta(1-p_o)\int_0^\infty tf_c(t)dt = (1-p_o)\frac{\eta}{\mu} = \frac{\eta}{\mu} \quad (\text{if } p_o = 0). \end{aligned}$$

Next, we show how to find the handoff call arrival rate from the handoff rate. It is easy to observe that for each nonblocking new call, there will be on the average $E[H]$ number of handoff calls induced, so the handoff call traffic will have arrival rate $\lambda_h = \lambda E[H]$. From Theorem 2.4, we obtain

Theorem 2.5. If the density function of the iid calling holding times has only finite possible isolated singular points, then the handoff call traffic has the arrival rate given by

$$\lambda_h = -\eta(1-p_o)\lambda\sum_{p\in\sigma_c}\operatorname*{Res}_{s=p}\frac{1-f^*(s)}{s^2[1-(1-p_f)f^*(s)]}f_c^*(-s). \tag{2.27}$$

The traffic intensity in a cell is given by

$$\rho = \lambda E[t_{no}] + \lambda_h E[t_{ho}], \tag{2.28}$$

where $E[t_{no}]$ and $E[t_{ho}]$ are the channel occupancy times for new calls and handoff calls, respectively.

Proof: We only need to prove the formula for the cell traffic intensity. The overall traffic arrival rate is $\lambda + \lambda_h$ and the expected channel occupancy time is

$$\frac{\lambda}{\lambda + \lambda_h} E[t_{no}] + \frac{\lambda_h}{\lambda + \lambda_h} E[t_{ho}],$$

hence (2.28) is straightforward. This completes the proof.

If the call holding times are exponentially distributed, then from Theorem 2.5 and (2.23), we have

$$\lambda_h = \frac{\eta(1 - p_o)[1 - f^*(\mu)]\lambda}{\mu[1 - (1 - p_f)f^*(\mu)]}, \tag{2.29}$$

and if we further assume that the nonprioritized handoff scheme is used (i.e., $p_o = p_f$), then from Theorem 2.5 we have

$$\rho = \frac{\lambda}{\mu} \left\{ 1 - \frac{p_o \eta [1 - f^*(\mu)]}{\mu[1 - (1 - p_o) f^*(\mu)]} \right\}. \tag{2.30}$$

These have been obtained in [36] using a different approach.

If the call holding times are Erlang distributed according to (2.24), then we have

$$\lambda_h = \lambda\eta(-1)^{m+1}\alpha^m(1 - p_o)\frac{g^{(m-1)}(\alpha)}{(m-1)!}, \quad \alpha = m\mu. \tag{2.31}$$

The cell traffic intensity can be computed from this formula and the expected channel occupancy time.

When the call holding times are Erlang distributed, we have to compute the derivatives of $g(s)$ (in (2.25)) for the computation of handoff rate and handoff call arrival rate. However, the explicit expressions for the derivatives of $g(s)$ may be difficult. We therefore develop the following recursive algorithm for their computations. Let

$$h(s) = s^2[1 - (1 - p_f)f^*(s)].$$

Using the formula

$$(uv)^{(p)} = \sum_{i=0}^{p} \binom{p}{i} u^{(i)} v^{(p-i)}$$

we obtain

$$
\begin{aligned}
h^{(0)}(\alpha) &= \alpha^2[1-(1-p_f)f^*(\alpha)] \\
h^{(1)}(\alpha) &= -\alpha^2(1-p_f)f^{*(1)}(\alpha)+2\alpha[1-(1-p_f)f^*(\alpha)] \\
h^{(2)}(\alpha) &= -\alpha^2(1-p_f)f^{*(2)}(\alpha)-4\alpha(1-p_f)f^{*(1)}(\alpha)+2[1-(1-p_f)f^*(\alpha)] \\
h^{(p)}(\alpha) &= -\alpha^2(1-p_f)f^{*(p)}(\alpha)-2p\alpha(1-p_f)f^{*(p-1)}(\alpha) \\
&\quad -p(p-1)(1-p_f)f^{*(p-2)}(\alpha), p\geq 3
\end{aligned}
$$

Since we have $g(s)h(s)=1-f^*(s)$, differentiating both sides we obtain $(p>0)$

$$\sum_0^p \binom{p}{i} g^{(i)}(s)h^{(p-i)}(s) = -f^{*(p)}(s).$$

From this, we obtain the following recursive algorithm to compute $g^{(m-1)}(\alpha)$:

$$
\begin{aligned}
g^{(0)}(\alpha) &= \frac{1-f^*(\alpha)}{h(\alpha)} \\
g^{(p)}(\alpha) &= -\frac{f^{*(p)}(\alpha)+\sum_{i=0}^{p-1}\binom{p}{i}g^{(i)}(\alpha)h^{(p-i)}(\alpha)}{h(\alpha)} \quad (p>0).
\end{aligned}
$$

2.5 Illustrative Examples

This section presents several demonstrative examples showing how the derived models can be used to study properties of channel occupancy times and handoff rate in the emerging PCS networks.

First, we assume that the call holding times are exponentially distributed with parameter μ. The following identity is useful for this example:

$$\int_t^\infty (\alpha x)^i e^{-\alpha x}dx = \frac{i!}{\alpha}\left(\sum_{k=0}^{i}\frac{(\alpha t)^k}{k!}\right)e^{-\alpha t}.$$

The distribution function $F(t)$ of the Erlang distribution is given by

$$F(t)=\Pr(t_m\leq t)=\int_0^t f(\tau)d\tau = 1-\sum_{i=0}^{m-1}\frac{(\alpha t)^i}{i!}e^{-\alpha t}.$$

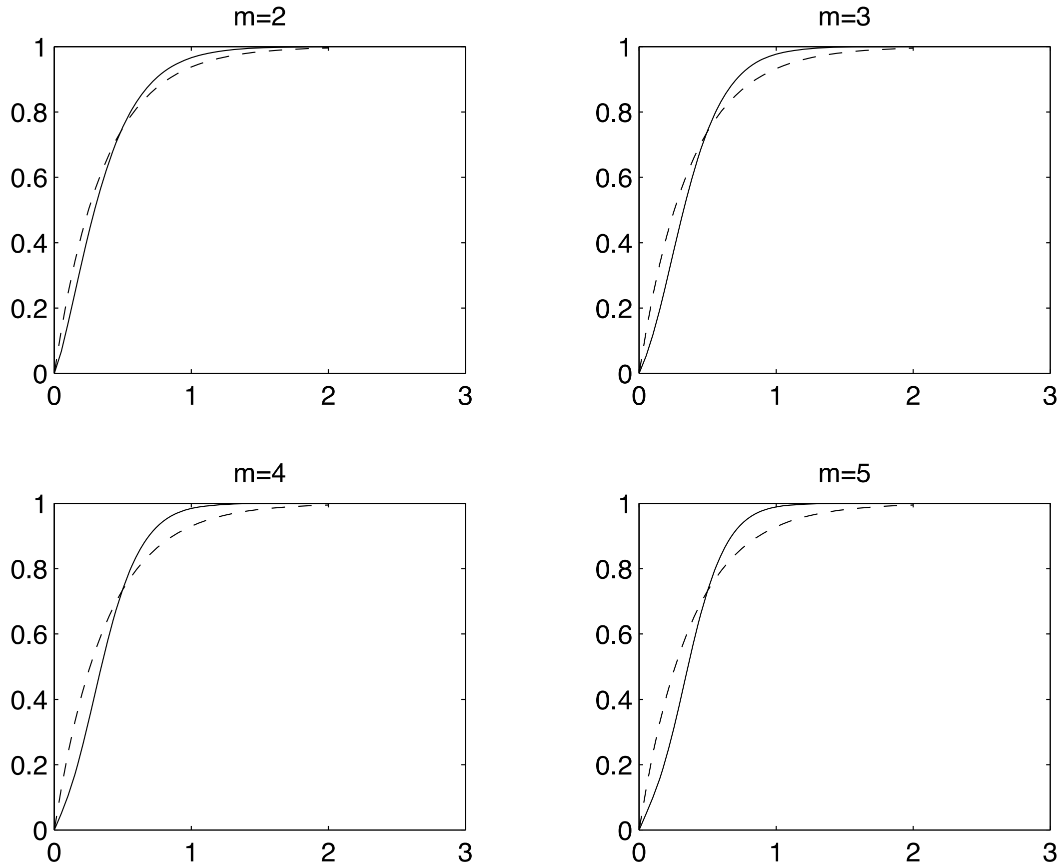


Figure 2.3: Distribution of handoff call channel occupancy time (solid line) and its exponential fitting (dashed line)

We consider the handoff call channel occupancy time distribution first. From (2.3), we can find the distribution function $F_{ho}(t)$ of the handoff call channel occupancy time as follows

$$
\begin{aligned}
F_{ho}(t) &= \Pr(t_{ho} \leq t) = \Pr(t_c \leq t) + [1 - \Pr(t_c \leq t)]F(t) \\
&= 1 - e^{-\mu t} + e^{-\mu t}\left[1 - \sum_{i=0}^{m-1} \frac{(\alpha t)^i}{i!} e^{-\alpha t}\right] \\
&= 1 - \left(\sum_{i=0}^{m-1} \frac{(\alpha t)^i}{i!}\right) e^{-(\alpha+\mu)t} \qquad (2.32)
\end{aligned}
$$

It is important to compare the real distribution function of handoff call channel occupancy time and its exponential fitting. Since expected value must be the same, we use the formula (2.13) to compute the expected value, and use its inverse as the parameter which determines the exponential function. Varying the shape parameter m from 2 to 5, we obtain the comparative plots in Figure 2.3. From this figure, we observe that the exponential approximation is not good, and it becomes worse when m grows. As we know, the variance of the cell residence time is $1/(m\eta^2)$, which decreases when m increases. Hence, we can conclude that the exponential approximation is not suitable for the handoff call channel occupancy time distribution when the variance of the cell residence time is very small.

Next, consider the new call channel occupancy times. Let $F_{no}(t)$ denote the distribution function of the new call channel occupancy time, then from (2.1), we obtain

$$\begin{aligned}
F_{no}(t) &= \Pr(t_c \le t) + [1 - \Pr(t_c \le t)] \Pr(r_1 \le t) = 1 - e^{-\mu t} + e^{-\mu t} \int_0^t f_r(\tau) d\tau \\
&= 1 - e^{\mu t} + e^{-\mu t} \left[1 - \int_t^\infty f_r(\tau) d\tau \right] = 1 - e^{-\mu t} \int_t^\infty f_r(\tau) d\tau \\
&= 1 - \eta e^{-\mu t} \int_t^\infty [1 - F(\tau)] d\tau = 1 - \eta e^{-\mu t} \int_t^\infty \sum_{i=0}^{m-1} \frac{(\alpha t)^i}{i!} e^{-\alpha \tau} d\tau \\
&= 1 - \eta e^{-\mu t} \sum_{i=0}^{m-1} \frac{1}{\alpha} \left(\sum_{k=0}^{i} \frac{(\alpha t)^k}{k!} \right) e^{-\alpha t} \\
&= 1 - \frac{1}{m} \sum_{i=0}^{m-1} (m - i) \frac{(\alpha t)^i}{i!} e^{-(\mu+\alpha)t} \quad \text{(noticing that } \alpha = m\mu) \\
&= 1 - \sum_{i=0}^{m-1} \left(1 - \frac{i}{m} \right) \frac{(\alpha t)^i}{i!} e^{-(\mu+\alpha)t}. \qquad (2.33)
\end{aligned}$$

From Figure 2.4, we observe that the new call channel occupancy time is less sensitive to the variance of cell residence times than the handoff call channel occupancy time, and the exponential fitting for new call channel occupancy time is better than the handoff call channel occupancy time.

From (2.32), (2.33) and Theorem 2.2, we can obtain the distribution $F_{co}(t)$ of channel occupancy time as follows

$$F_{co}(t) = \frac{\lambda}{\lambda + \lambda_h} F_{no}(t) + \frac{\lambda_h}{\lambda + \lambda_h} F_{ho}(t)$$

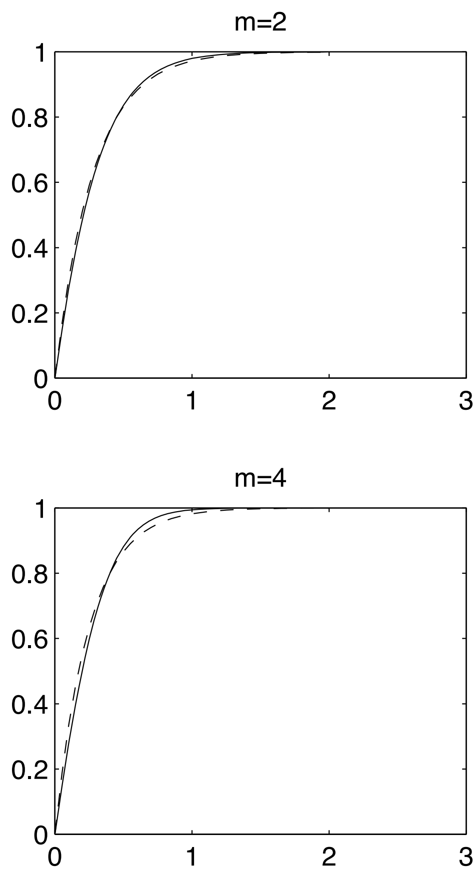


Figure 2.4: Distribution of new call channel occupancy time (solid line) and its exponential fitting (dashed line)

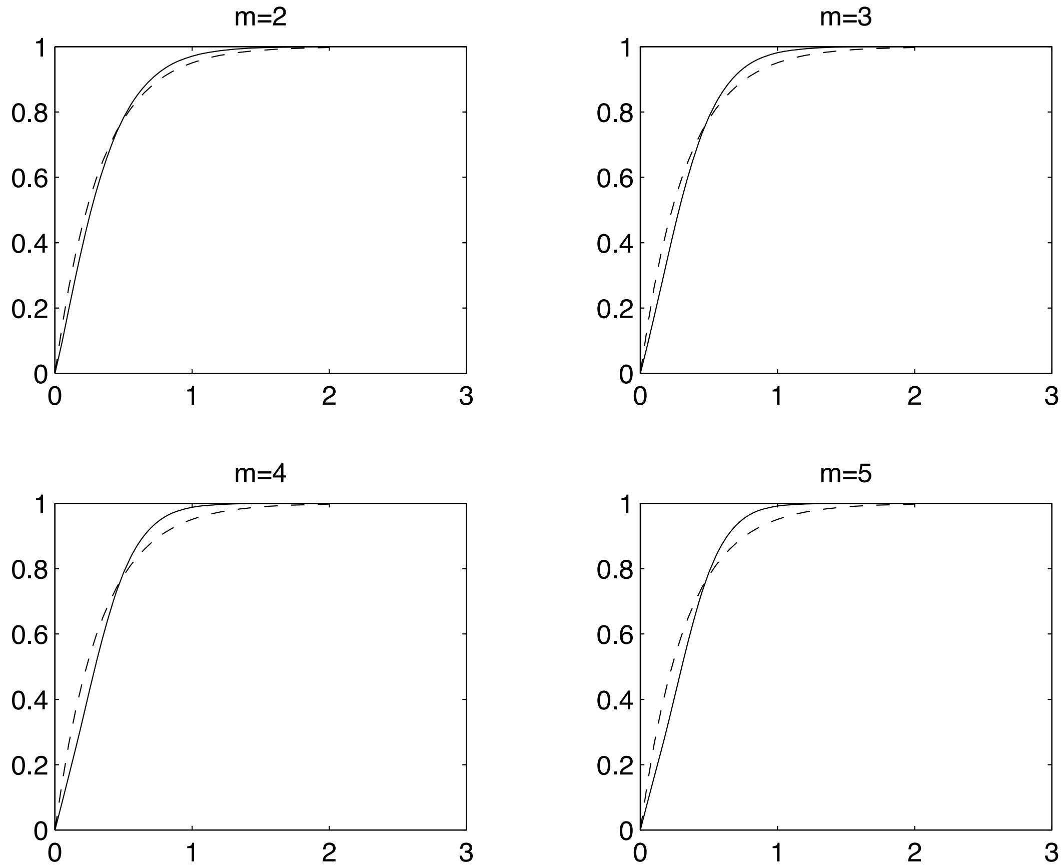


Figure 2.5: Distribution of channel occupancy time (solid line) and its exponential fitting (dashed line)

$$= \quad 1 - \sum_{i=0}^{m-1} \left(1 - \frac{\lambda i}{m(\lambda + \lambda_h)}\right) \frac{(\alpha t)^i}{i!} e^{-(\mu+\alpha)t}. \qquad (2.34)$$

Figure 2.5 shows the channel occupancy time (used in calculating the blocking probabilities). We observe that the channel occupancy time can be appropriately approximated by the exponential distribution when $m = 2$. However, there are significant discrepancies between the distributions of the actual channel occupancy time and the exponential distribution approximations when m becomes larger (i.e., the variance of the cell residence times becomes small).

Figures 2.6, 2.7 and 2.8 show the distributions of the new call channel occupancy time, handoff call channel occupancy time and channel occupancy time, respectively,

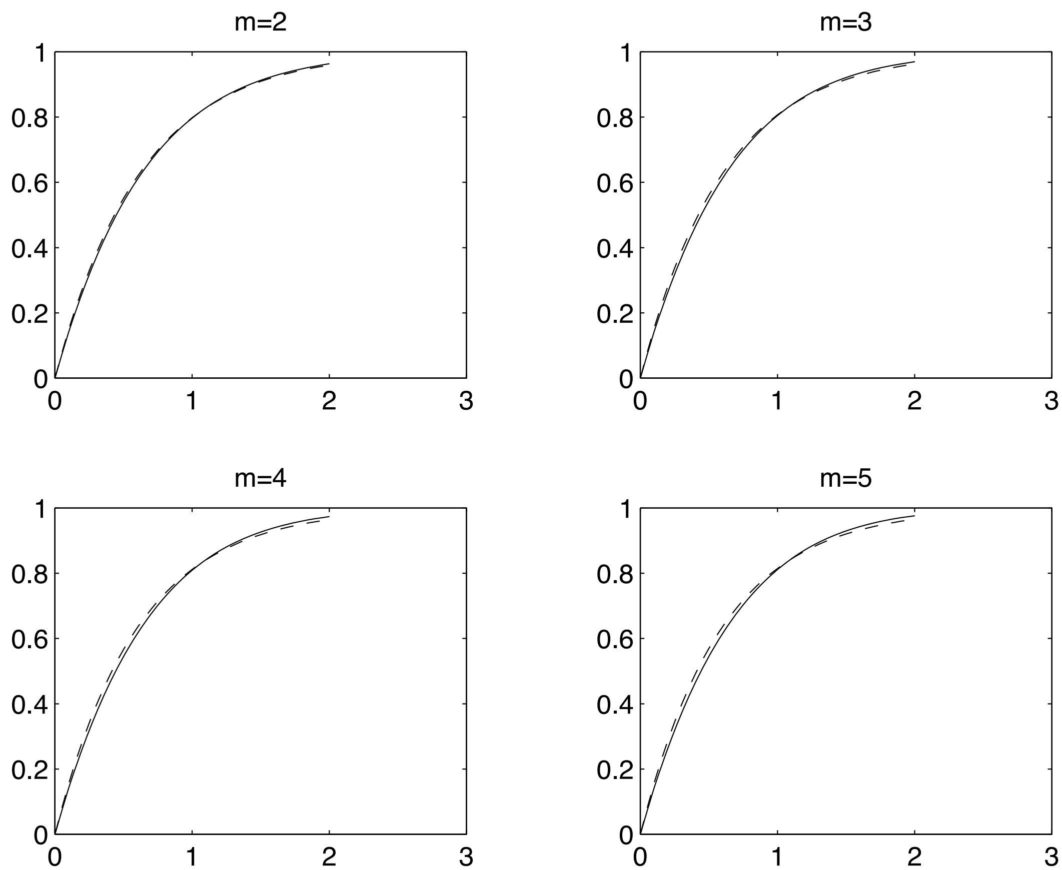


Figure 2.6: Distribution of new call channel occupancy time (solid line) and its exponential fitting (dashed line): small mobility η/μ

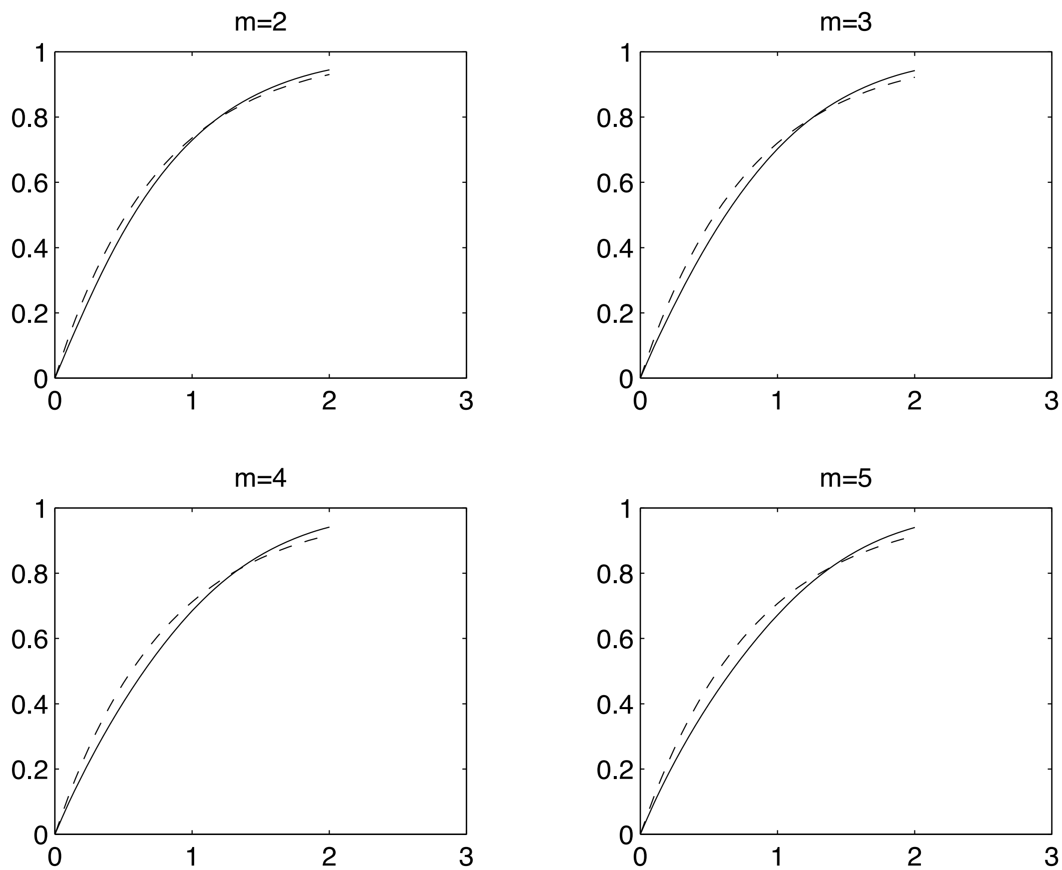


Figure 2.7: Distribution of handoff call channel occupancy time (solid line) and its exponential fitting (dashed line): small mobility η/μ

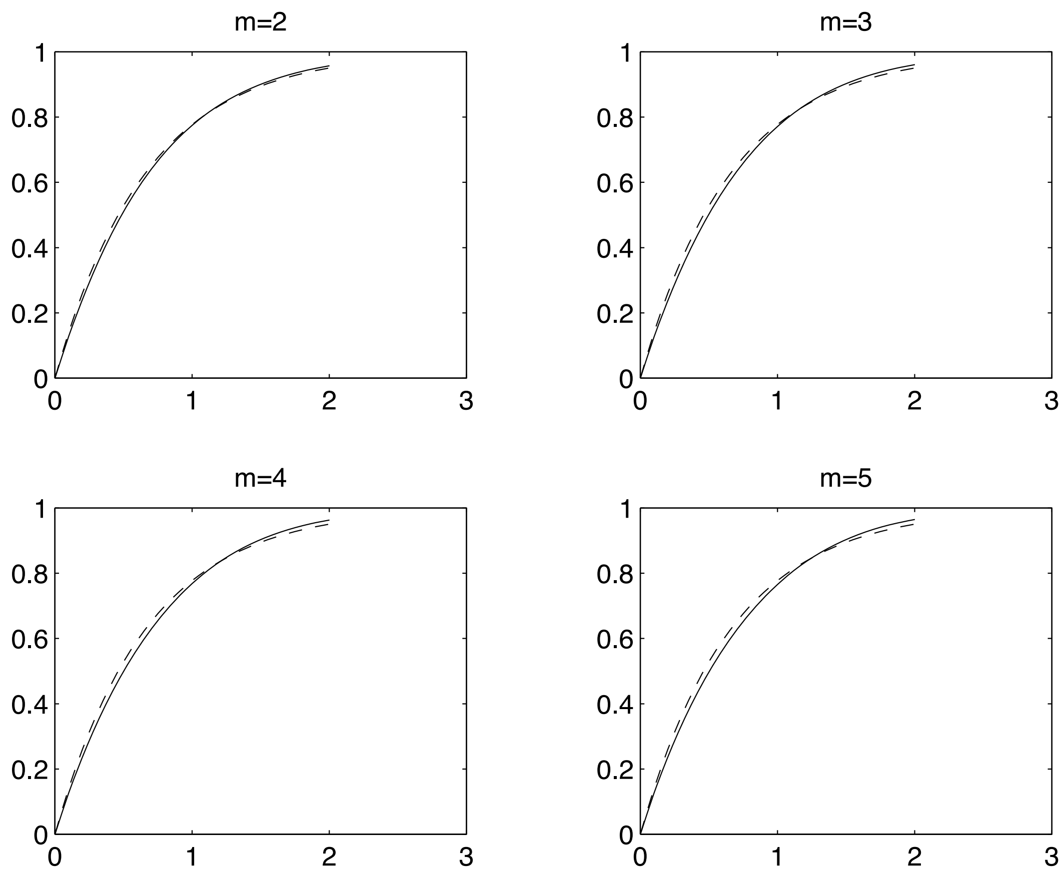


Figure 2.8: Distribution of channel occupancy time (solid line) and its exponential fitting (dashed line): small mobility η/μ

when the mobility η/μ is small (i.e., the customers are less mobile than the previous case). We observe that the new call channel occupancy time distribution has a good approximation by the exponential distribution, while there is still a significant mismatch between the handoff call channel occupancy time distribution and the exponential distribution. However, the channel occupancy time distribution can be better approximated by the exponential distribution.

Finally, we turn our attention to the expected number of handoffs. Assume that the call holding times are Erlang distributed. We shall use equation (2.26) and the recursive algorithms given at the end of the previous section. The cell residence times are Gamma distributed with the density function

$$f(t) = \frac{\beta^\gamma t^{\gamma-1}}{\Gamma(\gamma)} e^{-\beta t}, \; f^*(s) = \left(\frac{\beta}{s+\beta}\right)^\gamma, \; \beta = \gamma\eta,$$

where γ is the shape parameter, β is the scale parameter and the $\Gamma(\gamma)$ is the Gamma function. The mean and variance of this distribution are $1/\eta$ and $1/(\gamma\eta^2)$, respectively.

Figures 2.9 and 2.10 show the expected number of handoffs for different blocking probabilities and forced termination probabilities. We observe the following:

(a) The expected number of handoffs for fixed variance of the cell residence times (i.e., fixed m) is increasing as the mobility η/μ increases, as expected intuitively;

(b) The expected number of handoffs is increasing for fixed mobility as the variance of the cell residence times decreases (i.e., m is increasing);

(c) The expected number of handoffs is smaller than in the ideal case (when there is no blocking and no forced termination), which confirms our earlier expectations and conforms with general intuition;

(d) The expected number of handoffs is insensitive to the variance of the cell residence times when mobility η/μ is small;

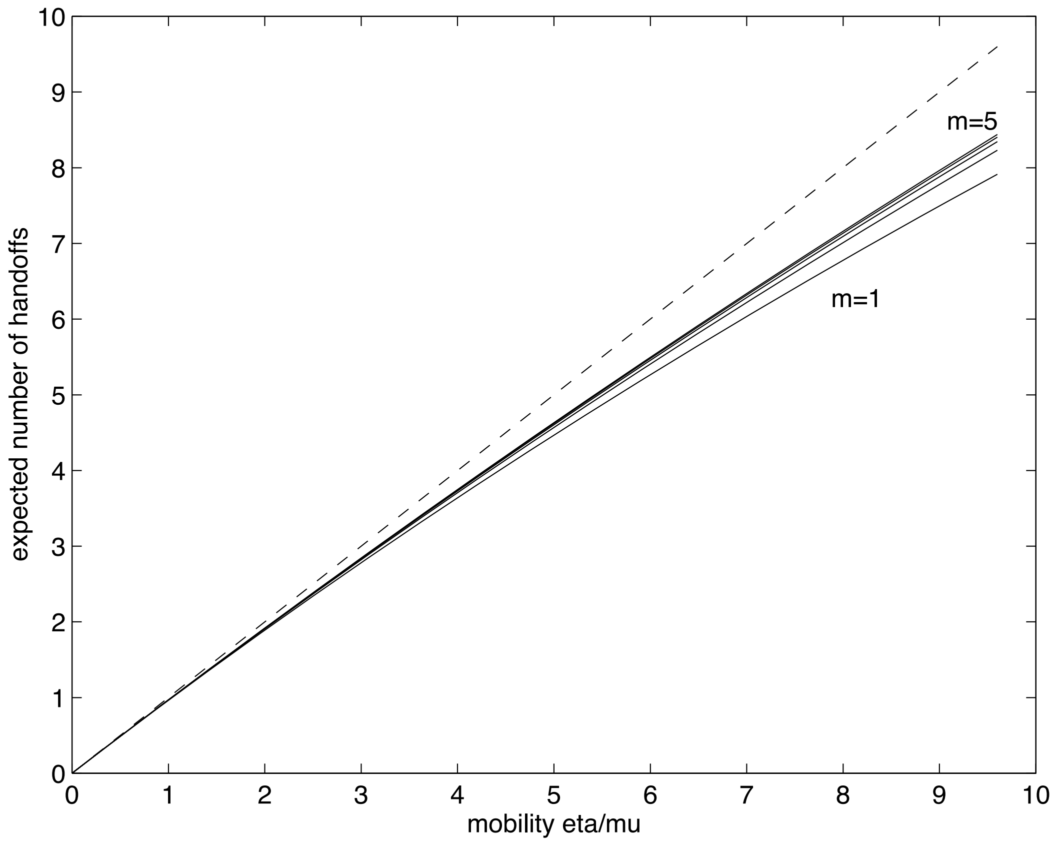


Figure 2.9: Expected number of handoffs: solid lines for $m = 1$–5 with $p_f = p_o = 0.02$, dashed line for the ideal case ($p_o = p_f = 0$)

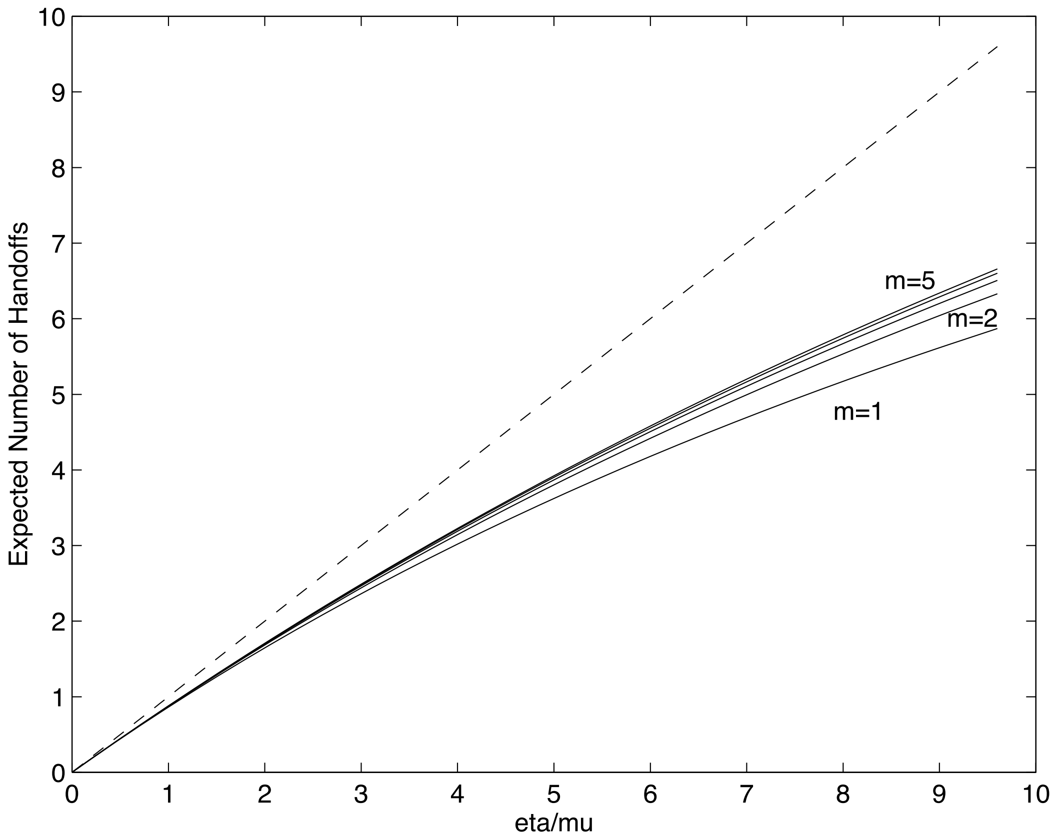


Figure 2.10: Expected number of handoffs: solid lines for $m = 1$–5 with $p_o = 0.1$ and $p_f = 0.05$, dashed line for the ideal case ($p_o = p_f = 0$)

(e) When handoff calls are given priority over new calls, the variance of the cell residence times affects the number of handoffs more significantly.

2.6 Conclusions

In this chapter, we have investigated the channel occupancy times and handoff rates (the expected number of handoffs) for GSM based mobile computing networks and integrated PCS networks. It has been shown that except for the case of exponentially distributed cell residence times, the channel occupancy time is not exponentially distributed, the handoff traffic is not Poisson, and the merged call traffic to a cell is not Poisson, contrary to commonly made assumptions in the past. We have derived the analytical expressions for channel occupancy time distributions, and provided comparisons with the real channel occupancy time distributions and the exponential fitting. The results presented here can be used to evaluate the goodness of the exponential distribution approximation. We further studied the handoff rate , and provided analytical results for generally distributed call holding times and associated cell residence times. The above results were demonstrated to be useful in the traffic study, performance evaluation, management and billing of PCS networks.

Chapter 3

CALL COMPLETION PROBABILITY AND EFFECTIVE CALL HOLDING TIMES

3.1 Introduction

In the last chapter, we have studied the cell traffic, channel occupancy times and handoff rate under very general assumptions. In this chapter, we study the second-level modeling: given the requirement for the blocking probability and the handoff failure probability, determine the call completion probability, the call incompletion probability and the effective call holding times under exponentially or Erlang distributed call holding time distribution and general cell residence time distribution.

In a PCS network, when a new call is originated and attempted in a cell, one of the channels assigned to the base station is used for the communication between the mobile portable and the base station as long as a channel is available. When all channels in a cell are in use while a new call (or handover call) is attempted in the cell, the call will be blocked and cleared from the system. When a call gets a channel, it will keep the channel until its completion, or until the mobile moves out of the cell, in which case the channel will be released for other use. When the mobile moves into a new cell while its call is ongoing, a new channel needs to be acquired in the new cell for further communication, using a handover procedure. During handover, if there is no channel available in the new cell for the "ongoing " call, it is forced to terminate before its completion ([29, 34]). In order to evaluate this behavior in a PCS network, the following three possibilities need to be considered.

- The call is blocked at its call initiation and is never connected (a *blocked new call*);
- The call is connected, successfully makes one or more handovers, but is forced to terminate before its completion because of the lack of an available channel (an *incomplete call*);
- The call is connected and completed (a *complete call*).

In a PCS system, one ideally wishes to allow all calls to be completed. Clearly, in the presence of channel availability limitations, this objective is not obtainable at all times. As an alternative, one wishes to make the call completion probability as high as the grade of service (GoS) needed.

When defining the right objective a major practical consideration is obviously pricing. The use of the same flat rate for both complete and incomplete calls is unfair and commercially unattractive. To stay competitive, a PCS network provider may apply discounts for calls that were not completed ([39]). However, since incomplete calls may spend significant time using PCS network resources, it is also impossible for the provider to apply a constant discount rate to all incomplete calls. In order to determine a reasonable (if not optimal) discount factor, it is therefore necessary to know for how long a call (either complete or incomplete) has used the network. The duration of the requested call connection is referred to as the *call holding time*. The duration of a actual call connection for an incomplete call will be called *the effective call holding time of an incomplete call* while the duration of a actual call connection of a complete call will be called *the effective call holding time of a complete call*. It is obvious that the duration of a requested call does not depend on the PCS network, it only depends on the mobile user (how long he wishes to maintain the call). For simplicity, we will term this duration the *call holding time*, as this is consistent with the ideal case when there are infinite numbers of channels and the handover procedure does not affect the duration of a connection. In reality, the number of radio channels is limited and handover procedure does come into play. The duration of a actual call connection will depend on PCS network: its traffic situation, channel availability, etc.

As we mentioned earlier, the performance modeling of a PCS network can be conducted at two levels. The first level modeling uses the number of radio channels in cells as an input parameter to determine the new call blocking probability and the forced termination probability. The second level modeling uses the new call blocking and the forced termination probabilities as input parameters to study the call completion probability (or the probability that a call is successfully complete) and the expected effective call holding times of a complete or an incomplete call. This chapter deals with second level modeling. The extension of our results to the first level modeling is possible and will be treated in the future. Since existing cellular systems are typically engineered at $1-2\%$ new call blocking and forced termination, these default values may be used as the reference input parameters for second level modeling. However, the call completion probability and the expected effective call holding times can not be derived directly from these two parameters. Both the cell residence time and the call holding time distributions must carefully be chosen to reflect the real systems. In our model, a general cell residence time distribution is considered, which can be used to accommodate any real PCS system. The selection of the call holding times were in the past typically assumed to be exponentially distributed. Such an assumption may be reasonable when the calls are charged based on the lengths of the call holding times. The assumption is no longer valid however for the modern telephone services which apply flat rate billing programs. Flat rate billing encourages people to make long calls (for example, people may log on from their PCs at home to main computers in the companies through local telephone calls and keep the connections for sometimes several days). Thus, with current communication environments long tails are observed for the call holding time distributions. In recent telephone network engineering ([3]), lognormal distributions ([22]) have been used to approximate the wireline call holding times. For most existing cellular systems, the wireless calls are charged based on the call holding times, and these systems can be appropriately modeled with the exponential call holding time distribution ([40, 43]). However, for the future PCS systems (especially the low power PCS systems such as CT-2 ([64]), DECT ([9]), or PACS ([52]), flat rate billing programs have been proposed. Thus, it is very important that we follow the wireline telephone network engineering approach; that is, use more general distribution to represent call holding time distribution.

In order to determine the call completion probability and the effective call holding times, we need to know the call arrival distribution, call holding time distribution and cell residence time (i.e. a mobile stays in a cell) (reflecting the frequency of handovers or the mobility of a mobile portable). We will call the duration of a call in a given cell the *cell residence time*. In this chapter, we will use the following assumptions:

- The call arrivals form a Poisson process;
- The cell residence times (the intervals that a portable stays in the cells) are independent, identically distributed (iid) non-lattice distribution;
- The call holding times are independent, identically distributed (iid) with non-lattice distribution.

The last two assumptions are a generalization of conventional analysis assumptions (e.g., the commonly used exponential distribution satisfies these two assumptions) and are chosen to fit the emerging PCS networks. The cell residence times should be general, since the time the mobile may spend in a cell may be highly variable due to the mobile speed and motion, and the geographic situation in the cell. For example in an urban area the routes the mobile users take may be highly irregular. The rate of the residence times is used to describe the mobility of the mobile portable.

In the traffic models studied in the literature, the call holding times are assumed to be exponentially distributed for reasons of tractability ([24, 68]). Under the assumption of exponentially distributed call holding times, Lin et al ([40, 43]) studied the performance of channel assignment strategies and obtained analytical results for forced termination probability and new call blocking probability. Under the same assumption, Lin and Chlamtac ([39]) analyzed the call completion probability and the expected effective call holding times.

In this chapter, we study the case of PCS networks using the generalized assumptions above. For this generalized model we obtain formulae for the call completion probability and the distribution (its Laplace transforms) of the effective call holding times of both complete and incomplete calls, from which expected effective call holding times can be derived. When the call holding times are Erlang distributed, easily computable formulae are given and recursive algorithms are developed.

3.2 Call Completion Probability

In this section, we study the call completion probability. Previously, Lin and Chlamtac ([39]) obtained a formula for the call completion probability for a PCS network with a general residence time distribution and exponential call holding time distribution. Here, we generalize these results to the case when call holding times

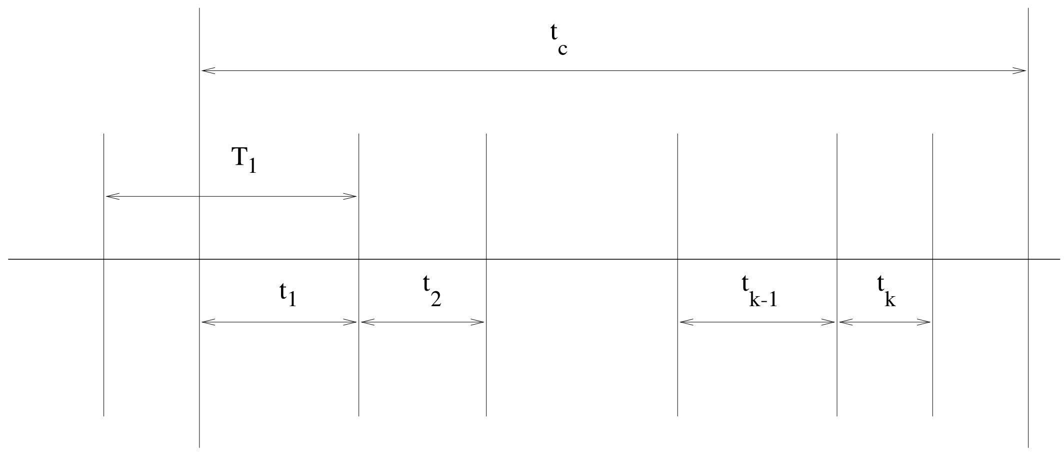


Figure 3.1: The timing diagram for a forced terminated call at kth handover

have a more general distribution.

We first consider the effective call holding time t for an incomplete call. Figure 3.1 illustrates the timing diagram for the call holding time, where T_1 is the time that the portable resides at cell 1, and t_i $(i \geq 2)$ is the residence time at cell i. According to our assumptions, $T_1, t_2, \ldots, t_k, \ldots$ are iid. Let t_i have non-lattice density function $f(\cdot)$ with the mean $1/\eta$ and $f^*(s)$ be the Laplace transform of $f(\cdot)$ (we will use $*$ to denote the Laplace transform following the tradition, [32]). Suppose that a call for the portable occurs when the portable is in cell 1. Let t_1 be the interval between the time instant when the call arrives and that when the portable moves out of cell 1. Let $r(t_1)$ and $r^*(s)$ be the density function and the Laplace transform of the t_1 distribution, respectively. From the renewal theory ([32]), t_1 is the residual life of the cell residence time of the portable in cell 1, so we have

$$r(t_1) = \eta \int_{t_1}^{\infty} f(\tau) d\tau \tag{3.1}$$

$$r^*(s) = \frac{\eta}{s}[1 - f^*(s)] \tag{3.2}$$

Consider the effective holding time $t = t_1 + t_2 + \cdots + t_k$ where $f_k(t)$ and $f_k^*(s)$ are its density function and Laplace transform. Since t_i $(i = 2, 3, \ldots, k)$ are iid, it is easy to derive

$$f_k^*(s) = E[\exp(-s(T_1 + \sum_{i=1}^{k} t_i))] = r^*(s)[f^*(s)]^{k-1} = \frac{\eta}{s}(1 - f^*(s))\,[f^*(s)]^{k-1}. \quad (3.3)$$

Let p_o be the probability that a new call attempt is blocked (i.e., the call is never connected), p_c be the probability that a call is completed (i.e., the call is connected and completed), p_f be the forced termination probability or the probability that no radio channel is available when a handoff call arrives. Then the probability of an incomplete call (i.e., the call is connected but is eventually forced to terminate) p_i is $1 - p_o - p_c$, which can be expressed as

$$\begin{aligned} 1 - p_o - p_c &= \sum_{k=1}^{\infty} \left\{ \int_0^\infty \int_t^\infty (1-p_o) f_k(t)(1-p_f)^{k-1} p_f f_c(t_c) dt_c dt \right\} \\ &= \sum_{k=1}^{\infty} \left\{ \int_0^\infty (1-p_o) f_k(t)(1-p_f)^{k-1} p_f \left[\int_t^\infty f_c(t_c) dt_c \right] dt \right\}, \quad (3.4) \end{aligned}$$

where $f_c(t_c)$ is the density function of the call holding times. In the first equation, the term in $\{\cdot\}$ is the probability that a call is forced to terminate at the kth handover (notice that the call is connected with probability $1 - p_o$, and then makes $k-1$ successful handovers with the probability $(1-p_f)^{k-1}$ and is forced to terminate at the kth handover with probability p_f). In the following derivation, we use the inverse Laplace transform formula. We will use σ to denote the real number with appropriate meaning used in the inverse Laplace transform formula ([35]). From (3.4), we have

$$\begin{aligned} & 1 - p_o - p_c = \\ = \; & (1-p_o)p_f \sum_{k=1}^{\infty} \int_0^\infty \left[\frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} f_k^*(s) e^{st} ds \right] (1-p_f)^{k-1} \left[\int_t^\infty f_c(t_c) dt_c \right] dt \\ = \; & \frac{(1-p_o)p_f}{2\pi j} \int_0^\infty \int_{\sigma - j\infty}^{\sigma + j\infty} \left[\sum_{k=1}^{\infty} f_k^*(s)(1-p_f)^{k-1} \right] e^{st} ds \left[\int_t^\infty f_c(t_c) dt_c \right] dt \\ = \; & \frac{(1-p_o)p_f}{2\pi j} \int_0^\infty \int_{\sigma - j\infty}^{\sigma + j\infty} \left[\sum_{k=1}^{\infty} \frac{\eta}{s}(1-f^*(s))[f^*(s)]^{k-1}(1-p_f)^{k-1} \right] e^{st} ds \\ & \quad \times \left[\int_t^\infty f_c(t_c) dt_c \right] dt \\ = \; & \frac{(1-p_o)p_f \eta}{2\pi j} \int_0^\infty \int_{\sigma - j\infty}^{\sigma + j\infty} \left[\sum_{k=1}^{\infty} \frac{1-f^*(s)}{s} [f^*(s)(1-p_f)]^{k-1} \right] e^{st} ds \end{aligned}$$

$$
\begin{aligned}
& \times \left[\int_t^\infty f_c(t_c)dt_c\right] dt \\
= & \frac{(1-p_o)p_f\eta}{2\pi j} \int_0^\infty \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f^*(s)}{s} \frac{1}{1-(1-p_f)f^*(s)} e^{st} ds \left[\int_t^\infty f_c(t_c)dt_c\right] dt \\
= & \frac{(1-p_o)p_f\eta}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]} \left\{\int_0^\infty e^{st} \left[\int_t^\infty f_c(t_c)dt_c\right] dt \right\} ds. \quad (3.5)
\end{aligned}
$$

Assume that the $\sigma_1 > 0$ is such that $e^{\sigma_1 t} \int_t^\infty f_c(t_c)dt_c \le M < \infty$ for any t (this will be true for most interesting distributions such as exponential or Erlang distribution), then choosing $\sigma < \sigma_1$, we have

$$
\begin{aligned}
& \int_0^\infty \left|e^{st} \int_t^\infty f_c(t_c)dt_c\right| dt = \int_0^\infty e^{\sigma t} \left|\int_t^\infty f_c(t_c)dt_c\right| dt \\
= & \int_0^\infty e^{-(\sigma_1-\sigma)t} \left|e^{\sigma t} \int_t^\infty f_c(t_c)dt_c\right| dt \le M \int_0^\infty e^{-(\sigma_1-\sigma)t} dt = M/(\sigma_1-\sigma) < \infty.
\end{aligned}
$$

Moreover, we have

$$
\begin{aligned}
& \int_0^\infty e^{st} \int_t^\infty f_c(t_c)dt_c dt \\
= & \frac{1}{s} e^{st} \int_t^\infty f_c(t_c)dt_c \Big|_0^\infty - \frac{1}{s} \int_0^\infty e^{st} \frac{d}{dt} \int_t^\infty f_c(t_c)dt_c dt \\
= & -\frac{1}{s} + \frac{1}{s} \int_0^\infty e^{st} f_c(t)dt = \frac{f^*(-s)-1}{s}.
\end{aligned}
$$

Taking this into (3.5), we obtain

$$
\begin{aligned}
1-p_o-p_c &= \frac{\eta(1-p_o)p_f}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]} \cdot \frac{f_c^*(-s)-1}{s} ds \\
&= -\eta(1-p_o)p_f \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{(1-f^*(s)(1-f_c^*(-s))}{s^2[1-(1-p_f)f^*(s)]} ds. \qquad (3.6)
\end{aligned}
$$

Since $f^*(s)$ has no poles in the right half complex plane, $|(1-p_f)f^*(s)| < 1$. Let σ_c denote the set of poles of $f_c^*(-s)$ in the right half complex plane (i.e., $-\sigma_c = \{-z | z \in \sigma_c\}$ is the set of poles of $f_c^*(s)$ in the left half plane). Choosing the contour that including the vertical line $z = \sigma$ and the semicircle $C_R = \{z = \sigma + Re^{j\theta} | -\pi/2 \le \theta \le \pi/2\}$, noticing that the integrand in (3.6) in the order of $1/R^2$ as $R \to \infty$, using the Residue Theorem ([35]), from (3.6) we obtain

$$
1-p_o-p_c = -\eta(1-p_o)p_f \left[- \operatorname*{Res}_{\substack{s=p \\ p\in\sigma_c}} \frac{(1-f^*(s))(1-f_c^*(-s))}{s^2[1-(1-p_f)f^*(s)]}\right]
$$

from which and the fact that the residue of an analytic function is zero we finally arrive at

Theorem 3.1. The probability that a call is completed is given by

$$p_c = (1 - p_o)\left[1 + \eta p_f \operatorname*{Res}_{\substack{s=p \\ p\in\sigma_c}} \frac{(1 - f^*(s))f_c^*(-s)}{s^2[1 - (1 - p_f)f^*(s)]}\right], \tag{3.7}$$

where $\operatorname*{Res}_{s=p}$ denotes the residue at pole $s = p$.

Remark: Since $f^*(s)$ is analytic and bounded by unity in the right half complex plane, the $(1-f^*(s))/s^2[1-(1-p_f)f^*(s)]$ is analytic in the right half plane, hence the function $(1 - f^*(s))f_c^*(-s)/s^2[1 - (1 - p_f)f^*(s)]$ has the same poles (including multiplicities) as the function $f_c^*(-s)$ in the right half complex plane. In the computation of the residue in Theorem 3.1, we only need to consider the poles and their corresponding multiplicities.

Next, we show how to compute p_c for a few specific cases. If the call holding times are exponentially distributed with the density function $f_c(t) = \mu e^{-\mu t}$, then its Laplace transform is given by $f_c^*(s) = \mu/(s + \mu)$ which has a simple pole μ in the left half plane. From Theorem 3.1, we obtain

$$\begin{aligned}
p_c &= (1 - p_o)\left[1 + \eta p_f \operatorname*{Res}_{\substack{s=p \\ p\in\sigma_c}} \frac{(1 - f^*(s))f_c^*(-s)}{s^2[1 - (1 - p_f)f^*(s)]}\right] \\
&= (1 - p_o)\left[1 + \eta p_f \operatorname*{Res}_{s=\mu} \frac{(1 - f^*(s))\mu/(-s + \mu)}{s^2[1 - (1 - p_f)f^*(s)]}\right] \\
&= (1 - p_o)\left[1 + \lim_{s\to\mu}(s - \mu)\frac{(1 - f^*(s))\mu/(-s + \mu)}{s^2[1 - (1 - p_f)f^*(s)]}\right] \\
&= (1 - p_o)\left[1 - \frac{\eta p_f(1 - f^*(\mu))}{\mu[1 - (1 - p_f)f^*(\mu)]}\right].
\end{aligned} \tag{3.8}$$

This is obtained in [39] using a different approach.

Assume now that the call holding times are Erlang distributed with the following density function

$$f_c(t_c) = \frac{(\alpha t_c)^{m-1}}{(m - 1)!}\alpha e^{-\alpha t_c}, \ m = 1, 2, \ldots$$

where m is the shape parameter and $\alpha = m\mu$ is the scale parameter. This density function has the following Laplace transform

$$f_c^*(s) = \left(\frac{\alpha}{s+\alpha}\right)^m$$

and $f_c^*(-s)$ has a unique pole at α in the right half complex plane. Let

$$g(s) = \frac{1-f^*(s)}{s^2[1-(1-p_f)f^*(s)]}.$$

As we remarked before, this function is analytic in the right half complex plane. From Theorem 3.1, we obtain

$$\begin{aligned}
p_c &= (1-p_o)\left[1+\eta p_f \operatorname*{Res}_{s=\alpha} g(s)\left(\frac{\alpha}{-s+\alpha}\right)^m\right] \\
&= (1-p_o)\left[1+\frac{\eta p_f}{(m-1)!}\lim_{s\to\alpha}\frac{d^{m-1}}{ds^{m-1}}\left((-\alpha)^m(s-\alpha)^m\frac{g(s)}{(s-\alpha)^m}\right)\right] \\
&= (1-p_o)[1+(-\alpha)^m\frac{\eta p_f}{(m-1)!}g^{(m-1)}(\alpha)].
\end{aligned}$$

Thus, we have

Corollary 3.1. For a PCS network with Erlang call holding times, the probability of a call completion is given by

$$p_c = (1-p_o)[1+(-\alpha)^m\frac{\eta p_f}{(m-1)!}g^{(m-1)}(\alpha)], \tag{3.9}$$

where $g^{(m-1)}(\alpha)$ denotes the derivative of $(m-1)$th order.

When $m=1$, we have

$$p_c = (1-p_o)[1-\eta p_f \alpha g(\alpha)] = (1-p_o)\left\{1-\frac{\eta p_f(1-f^*(\alpha))}{\alpha[1-(1-p_f)f^*(\alpha)]}\right\},$$

this is the same as (3.8) ($\alpha=\mu$ in this case), which is not surprising because Erlang distribution is exponential when $m=1$.

When $m=2$, since

$$\begin{aligned}
g'(s) &= \frac{d}{ds}\left(\frac{1-f^*(s)}{s^2[1-(1-p_f)f^*(s)]}\right) \\
&= -\frac{f^{*(1)}(s)}{s^2[1-(1-p_f)f^*(s)]}-\frac{2(1-f^*(s))}{s^3[1-(1-p_f)f^*(s)]}+\frac{(1-p_f)(1-f^*(s))f^{*(1)}(s)}{s^2[1-(1-p_f)f^*(s)]^2}.
\end{aligned}$$

Thus, from Corollary 3.1, we obtain

$$
\begin{aligned}
p_c &= (1-p_o)[1+\eta p_f \alpha^2 g'(\alpha)] \\
&= (1-p_o)\left\{1-\frac{\eta p_f(1-f^*(\alpha))}{\alpha[1-(1-p_f)f^*(\alpha)]}\right\} - \alpha(1-p_o)\eta p_f \\
&\times \left\{\frac{f^{*(1)}(\alpha)/\alpha+(1-f^*(\alpha))/\alpha^2}{1-(1-p_f)f^*(\alpha)} - \frac{(1-p_f)(1-f^*(\alpha))f^{*(1)}(\alpha)/\alpha}{[1-(1-p_f)f^*(\alpha)]^2}\right\}.
\end{aligned}
\tag{3.10}
$$

It is observed that as m increases, the computation becomes much involved. We will give a recursive algorithm for the computation of $g^{(m)}(\alpha)$ which is needed in Corollary 3.1. Let

$$h(s) = s^2[1-(1-p_f)f^*(s)]. \tag{3.11}$$

Using the formula

$$(uv)^{(p)} = \sum_{i=0}^{p}\binom{p}{i}u^{(i)}v^{(p-i)} \tag{3.12}$$

we obtain

$$
\begin{aligned}
h^{(0)}(\alpha) &= \alpha^2[1-(1-p_f)f^*(\alpha)] \\
h^{(1)}(\alpha) &= -\alpha^2(1-p_f)f^{*(1)}(\alpha)+2\alpha[1-(1-p_f)f^*(\alpha)] \\
h^{(2)}(\alpha) &= -\alpha^2(1-p_f)f^{*(2)}(\alpha)-4\alpha(1-p_f)f^{*(1)}(\alpha)+2[1-(1-p_f)f^*(\alpha)] \\
h^{(p)}(\alpha) &= -\alpha^2(1-p_f)f^{*(p)}(\alpha)-2p\alpha(1-p_f)f^{*(p-1)}(\alpha) \\
&\quad -p(p-1)(1-p_f)f^{*(p-2)}(\alpha),\ \ p\geq 3
\end{aligned}
\tag{3.13}
$$

Since we have $g(s)h(s) = 1-f^*(s)$, differentiating both sides and applying (3.12), we obtain ($p>0$)

$$\sum_{0}^{p}\binom{p}{i}g^{(i)}(s)h^{(p-i)}(s) = -f^{*(p)}(s).$$

From this, we obtain the following recursive algorithm to compute $g^{(m-1)}(\alpha)$:

$$
\begin{aligned}
g^{(0)}(\alpha) &= \frac{1-f^*(\alpha)}{h(\alpha)} \\
g^{(p)}(\alpha) &= -\frac{f^{*(p)}(\alpha)+\sum_{i=0}^{p-1}\binom{p}{i}g^{(i)}(\alpha)h^{(p-i)}(\alpha)}{h(\alpha)} \quad (p>0)
\end{aligned}
\tag{3.14}
$$

Thus, using (3.13) and (3.14), we can easily compute $g^{(m-1)}(\alpha)$, hence p_c from Corollary 3.1.

3.3 The Expected Effective Call Holding Times

In the preceding section we discussed the probability for a call to complete. However, this quantity does not address the time needed for a call to complete. It is desirable to know how much time is needed for a complete call to finish and how much time does an incomplete call spend using the resource (bandwidth) so that an appropriate pricing scheme can be devised. In this section, we present a solution to this problem.

Similar to the argument in [39], the density function for the effective call holding time of an incomplete call that is forced to terminate is given by

$$\begin{aligned} g_i(t) &= \left(\frac{1}{p_i}\right)\left[\sum_{k=1}^{\infty} f_k(t)(1-p_o)(1-p_f)^{k-1}p_f\int_t^{\infty} f_c(t_c)dt_c\right] \\ &= \left(\frac{1-p_o}{1-p_c-p_o}\right)\left[\sum_{k=1}^{\infty} f_k(t)(1-p_f)^{k-1}p_f\int_t^{\infty} f_c(t_c)dt_c\right] \qquad (3.15)\end{aligned}$$

where $p_i = 1 - p_o - p_c$ denotes the probability of a call to be incomplete and p_c is computed in the last section. In the first equation, the term under the summation is the density that the call is forced to terminate after k handovers.

Next, we want to find the Laplace transform of $g_i(z)$ from which the expected value can be easily obtained. From (3.15), we have

$$g_i^*(z) = \int_0^{\infty} e^{-zt}g_i(t)dt = \frac{(1-p_o)p_f}{p_i}\int_0^{\infty} e^{-zt}\sum_{k=1}^{\infty} f_k(t)(1-p_f)^{k-1}\int_t^{\infty} f_c(t_c)dt_c dt$$

$$\begin{aligned} &= \frac{(1-p_o)p_f}{p_i}\sum_{k=1}^{\infty}\int_0^{\infty}\left[\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} f_k^*(s)e^{st}ds\right]e^{-zt}(1-p_f)^{k-1}\left[\int_t^{\infty} f_c(t_c)dt_c\right]dt \\ &= \frac{(1-p_o)p_f}{2p_i\pi j}\int_0^{\infty}\int_{\sigma-j\infty}^{\sigma+j\infty}\left[\sum_{k=1}^{\infty} f_k^*(s)(1-p_f)^{k-1}\right]e^{(s-z)t}ds\left[\int_t^{\infty} f_c(t_c)dt_c\right]dt \\ &= \frac{(1-p_o)p_f}{2p_i\pi j}\int_0^{\infty}\int_{\sigma-j\infty}^{\sigma+j\infty}\left[\sum_{k=1}^{\infty}\frac{\eta}{s}(1-f^*(s))[f^*(s)]^{k-1}(1-p_f)^{k-1}\right]e^{(s-z)t}ds\end{aligned}$$

$$
\begin{aligned}
&\times \left[\int_t^\infty f_c(t_c)dt_c\right] dt \\
= \;& \frac{(1-p_o)p_f\eta}{2p_i\pi j}\int_0^\infty \int_{\sigma-j\infty}^{\sigma+j\infty}\left[\sum_{k=1}^{\infty}\frac{1-f^*(s)}{s}[f^*(s)(1-p_f)]^{k-1}\right]e^{(s-z)t}ds \\
&\times \left[\int_t^\infty f_c(t_c)dt_c\right] dt \\
= \;& \frac{(1-p_o)p_f\eta}{2p_i\pi j}\int_0^\infty \int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s}\frac{1}{1-(1-p_f)f^*(s)}e^{(s-z)t}ds\left[\int_t^\infty f_c(t_c)dt_c\right] dt \\
= \;& \frac{(1-p_o)p_f\eta}{2p_i\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}\left\{\int_0^\infty e^{(s-z)t}\left[\int_t^\infty f_c(t_c)dt_c\right] dt\right\} ds.
\end{aligned}
\tag{3.16}
$$

Similar to the argument in the last section, we have

$$
\int_0^\infty e^{(s-z)t}\int_t^\infty f_c(t_c)dt_c dt = -\frac{1}{s-z}+\frac{1}{s-z}\int_0^\infty e^{(s-z)t}f_c(t)dt = \frac{f^*(-s+z)-1}{s-z}.
$$

Taking this into (3.16), we obtain

$$
g_i^*(z) = \frac{\eta(1-p_o)p_f}{2p_i\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}\cdot\frac{f_c^*(-s+z)-1}{s-z}ds. \tag{3.17}
$$

Since

$$
\lim_{s\to z}\frac{f_c^*(-s+z)-1}{s-z} = f_c^{*(1)}(0),
$$

so $s = z$ is a removable singular point ([35]) of the integrand of (3.17). Thus, the poles of the integrand in the right half complex plane is those of $f_c^*(-s+z)$, i.e., $\{z+p|p\in\sigma_c\}$. As in the last section, we obtain

Theorem 3.2.

$$
g_i^*(z) = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{(1-f^*(s))(f_c^*(-s+z)-1)}{s(s-z)[1-(1-p_f)f^*(s)]}. \tag{3.18}
$$

Now, assume that the call holding times are Erlang distributed with parameter (m,α). Then $f_c^*(s) = (\alpha/(s+\alpha))^m$. Let

$$
g_1(s) = \frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}. \tag{3.19}
$$

Then, from Theorem 3.2, we have

$$\begin{aligned}
g_i^*(z) &= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \operatorname*{Res}_{s=z+\alpha} g_1(s) \left[\frac{\left(\frac{\alpha}{-s+z+\alpha}\right)^m - 1}{s-z} \right] \\
&= -(-1)^m \frac{\alpha^m \eta(1-p_o)p_f}{p_i} \operatorname*{Res}_{s=z+\alpha} \frac{g_1(s)}{(s-z)[s-(z+\alpha)]^m} \\
&= (-1)^{m-1} \frac{\alpha^m \eta(1-p_o)p_f}{(m-1)!p_i} \lim_{s\to z+\alpha} \frac{d^{m-1}}{ds^{m-1}} \left[(s-(z+\alpha))^m \frac{g_1(s)}{(s-z)[s-(z+\alpha)]^m} \right] \\
&= (-1)^{m-1} \frac{\alpha^m \eta(1-p_o)p_f}{(m-1)!p_i} \lim_{s\to z+\alpha} \frac{d^{m-1}}{ds^{m-1}} \left[\frac{1}{s-z} g_1(s) \right] \\
&= (-1)^{m-1} \frac{\alpha^m \eta(1-p_o)p_f}{(m-1)!p_i} \lim_{s\to z+\alpha} \sum_{j=0}^{m-1} \binom{m-1}{j} \left(\frac{1}{s-z} \right)^{(j)} g_1^{(m-1-j)}(s) \\
&= (-1)^{m-1} \frac{\alpha^m \eta(1-p_o)p_f}{(m-1)!p_i} \lim_{s\to z+\alpha} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(-1)^j j!}{(s-z)^{j+1}} g_1^{(m-1-j)}(s) \\
&= (-1)^{m-1} \frac{\alpha^m \eta(1-p_o)p_f}{(m-1)!p_i} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(-1)^j j!}{\alpha^{j+1}} g_1^{(m-1-j)}(z+\alpha) \\
&= \frac{\eta(1-p_o)p_f}{p_i} \sum_{j=0}^{m-1} \frac{(-\alpha)^{m-1-j}}{(m-1-j)!} g_1^{(m-1-j)}(z+\alpha) \\
&= \frac{\eta(1-p_o)p_f}{p_i} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j)}(z+\alpha).
\end{aligned} \tag{3.20}$$

Finally, we arrive at

Theorem 3.3. For a PCS network with Erlang distributed call holding times, the Laplace transform of the density function of the effective call holding time of an incomplete call is given by

$$g_i^*(z) = \frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j)}(z+\alpha), \tag{3.21}$$

the expected effective call holding time of an incomplete call is given by

$$T_i = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j+1)}(\alpha), \tag{3.22}$$

and the variance of the effective incomplete call holding times is given by

$$V_i = \frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j+2)}(\alpha) - \left[\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j+1)}(\alpha) \right]^2. \tag{3.23}$$

Proof: (3.22) can be obtained by noticing that $T_i = -g_i^{*(1)}(0)$. (3.23) can be proved by the following relationship: $Var(X) = E(X^2) - (E(X))^2$ for any random variable X.

Remark: In fact, from equation (3.21), we can easily find all moments of the effective call holding time of an incomplete call, which is easily given by (let $T_i(k)$ denote the kth moment)

$$T_i(k) = (-1)^k \frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{j=0}^{m-1} \frac{(-\alpha)^j}{j!} g_1^{(j+k)}(\alpha), \;\; k \geq 1.$$

When $m = 1$, the Erlang distribution is the exponential distribution. In this case, we have

$$g_i^*(z) = \frac{\eta(1-p_o)p_f}{1-p_o-p_c} g_1(z+\alpha),$$

and

$$\begin{aligned} T_i &= \frac{\eta(1-p_o)p_f}{1-p_o-p_c} g_1^{(1)}(0) = \frac{\eta(1-p_o)p_f}{\alpha(1-p_o-p_c)[1-(1-p_f)f^*(\alpha)]} \\ &\quad \times \left\{ \frac{1-f^*(\alpha)}{\alpha} + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)} \right\} \end{aligned} \tag{3.24}$$

This is the same as in [39], where a different approach is used. Using (3.8) in (3.24), we obtain

$$T_i = \frac{1}{\alpha} + \frac{p_f f^{*(1)}(\alpha)}{[1-f^*(\alpha)][1-(1-p_f)f^*(\alpha)]}.$$

Since $f^{*(1)}(\alpha) < 0$ and $0 \leq f^*(\alpha) \leq 1$, we have $T_i \leq 1/\alpha$, i.e., the expected effective call holding time of an incomplete call is less than the expected non-interrupted call holding time.

For $m > 1$, we can also give a recursive algorithm as before to compute $g_1^{(p)}(\alpha)$ which is needed in (3.22). Let

$$h_1(s) = s[1-(1-p_f)f^*(s)]. \tag{3.25}$$

Then, we have

$$h_1^{(0)}(\alpha) = \alpha[1-(1-p_f)f^*(\alpha)]$$

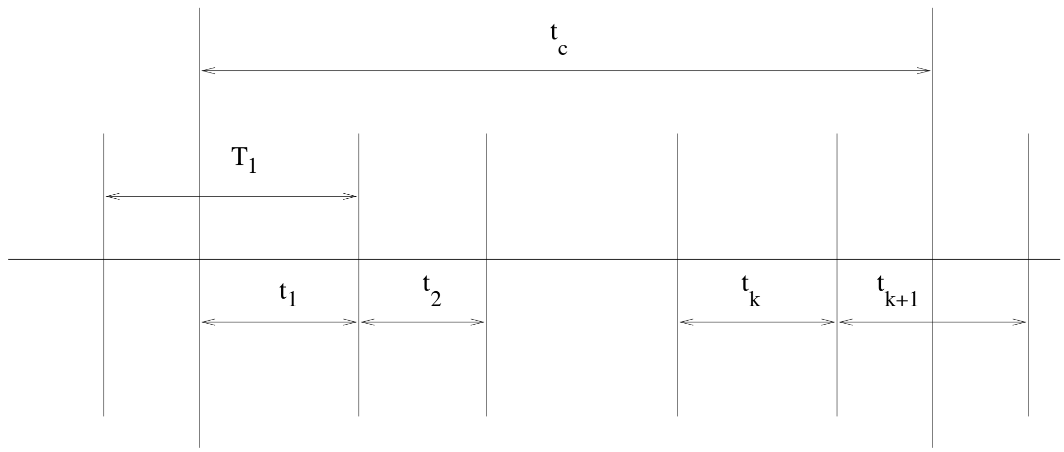


Figure 3.2: The timing diagram for the effective call times of a complete call: the call completes after k handovers

$$
\begin{aligned}
h_1^{(1)}(\alpha) &= -\alpha(1-p_f)f^{*(1)}(\alpha) + [1-(1-p_f)f^*(\alpha)] \\
h^{(p)}(\alpha) &= -\alpha(1-p_f)f^{*(p)}(\alpha) - p(1-p_f)f^{*(p-1)}(\alpha), \quad p \geq 2
\end{aligned} \tag{3.26}
$$

and

$$
\begin{aligned}
g_1^{(0)}(\alpha) &= \frac{1-f^*(\alpha)}{h_1(\alpha)} \\
g_1^{(p)}(\alpha) &= -\frac{f^{*(p)}(\alpha) + \sum_{i=0}^{p-1} \binom{p}{i} g_1^{(i)}(\alpha) h_1^{(p-i)}(\alpha)}{h_1(\alpha)} \quad (p > 0)
\end{aligned} \tag{3.27}
$$

Thus, using (3.26) and (3.27), we can easily compute $g_1^{(p-1)}(\alpha)$ $(p = 1, 2, \ldots, m)$, hence T_i from Theorem 3.3.

Next, we study the expected effective holding time for a complete call. The timing diagram is shown in Figure 3.2, in which the call is completed when the portable is in cell k'. t_c represents as before the effective call holding time for a complete call. If $k' = 1$, $0 \leq t_c \leq t_1$; while if $k' > 1$, $t_1 + t_2 + \cdots + t_{k'-1} \leq t_c \leq t_1 + t_2 + \cdots + t_{k'}$. Let $k = k' - 1$, then we have

$$
\text{For } k = 0,\ 0 \leq t_c \leq t_1 \tag{3.28}
$$

$$
\text{For } k > 0,\ t_1 + t_2 + \cdots + t_k \leq t_c \leq t_1 + t_2 + \cdots + t_{k+1} \tag{3.29}
$$

Using a similar argument in [39] (or a simple conditional probability argument), we can obtain the density function $g_c(t_c)$ of the effective call holding time of a complete call is given by

$$g_c(t_c) = U(t_c) + W(t_c) \tag{3.30}$$

where

$$\begin{aligned} U(t_c) &= \left(\frac{1-p_o}{p_c}\right)\left[f_c(t_c)\int_{t_c}^{\infty} r(t_1)dt_1\right], & (3.31)\\ W(t_c) &= \left(\frac{1-p_o}{p_c}\right)\left[\sum_{k=1}^{\infty} f_c(t_c)\int_0^{t_c}\int_{t_c-t}^{\infty} f_k(t)(1-p_f)^k f(\tau)d\tau dt\right]. & (3.32)\end{aligned}$$

$U(t_c)$ corresponds to (3.28) and $W(t_c)$ corresponds to (3.29), where $(1-p_o)$ is the probability of nonblocking, $(1-p_f)$ is the probability of no forced termination. The equation (3.30) can be derived from $P(t_c \le x) = \sum_{k=0}^{\infty} P(t_c \le x, k)$ where $P(t_c \le x, k)$ denotes the probability that the call is completed in cell $k+1$ and the effective call holding time is not exceeding x. Rigorous derivation can be obtained following a similar argument in [40].

We first find the Laplace transforms of $U(t_c)$ and $W(t_c)$. From (3.31) and the Residue Theorem, we have

$$\begin{aligned} U^*(z) &= \int_0^{\infty} e^{-zt_c}U(t_c)dt_c = \left(\frac{1-p_o}{p_c}\right)\int_0^{\infty} e^{-zt_c}f_c(t_c)\left(\int_{t_c}^{\infty} r(t_1)dt_1\right)dt_c \\ &= \left(\frac{1-p_o}{p_c}\right)\int_0^{\infty} e^{-zt_c}f_c(t_c)\frac{1}{2\pi j}\left(\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-r^*(s)}{s}e^{st_c}ds\right)dt_c \\ &= \left(\frac{1-p_o}{p_c}\right)\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-r^*(s)}{s}\left(\int_0^{\infty} f_c(t_c)e^{-(z-s)t_c}dt_c\right)ds \\ &= \left(\frac{1-p_o}{p_c}\right)\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{s-\eta(1-f^*(s))}{s^2}f_c^*(z-s)ds \\ &= -\left(\frac{1-p_o}{p_c}\right)\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{s-\eta(1-f^*(s))}{s^2}f_c^*(-s+z). \qquad (3.33)\end{aligned}$$

Since

$$\begin{aligned} \frac{d}{dt_c}\left[\int_0^{t_c} f_k(t)\int_{t_c-t}^{\infty} f(\tau)d\tau dt\right] &= f_k(t_c)\int_0^{\infty} f(\tau)d\tau - \int_0^{t_c} f_k(t)f(t_c-t)dt \\ &= f_k(t_c) - f_k(t_c) * f(t_c),\end{aligned}$$

where $*$ denotes the convolution operator, hence the Laplace transform of $\int_0^{t_c} f_k(t)\int_{t_c-t}^{\infty} f(\tau)d\tau dt_c$ is $(f_k^*(s) - f_k^*(s)f^*(s))/s = (1-f^*(s))f_k^*(s)/s$. From (3.32) and the Residue Theorem, we have

$$\begin{aligned}
W^*(z) &= \int_0^\infty e^{-zt_c} W(t_c)dt_c \\
&= \frac{1-p_o}{p_c}\sum_{k=1}^{\infty}(1-p_f)^k \int_0^\infty e^{-zt_c} f_c(t_c)\left[\int_0^{t_c} f_k(t)\int_{t_c-t}^{\infty} f(\tau)d\tau dt\right] dt_c \\
&= \frac{1-p_o}{p_c}\sum_{k=1}^{\infty}(1-p_f)^k e^{-zt_c} f_c(t_c)\left[\frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s} f_k^*(s)e^{st_c}ds\right] dt_c \\
&= \frac{(1-p_o)(1-p_f)}{2\pi p_c j}\int_0^\infty e^{-zt_c} f_c(t_c)\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s} \\
&\quad \times \left[\sum_{k=1}^{\infty} f_k^*(s)(1-p_f)^{k-1}\right] e^{st_c}dsdt_c \\
&= \frac{(1-p_o)(1-p_f)}{2\pi p_c j}\int_0^\infty e^{-zt_c} f_c(t_c)\int_{\sigma-j\infty}^{\sigma+j\infty}\eta\left[\frac{1-f^*(s)}{s}\right]^2 \\
&\quad \times \left[\sum_{k=1}^{\infty}[(1-p_f)f^*(s)]^{k-1}\right] e^{st_c}dsdt_c \\
&= \frac{\eta(1-p_o)(1-p_f)}{2\pi p_c j}\int_0^\infty e^{-zt_c} f_c(t_c)\int_{\sigma-j\infty}^{\sigma+j\infty}\left[\frac{1-f^*(s)}{s}\right]^2 \\
&\quad \times \frac{1}{1-(1-p_f)f^*(s)} e^{st_c}dsdt_c \\
&= \frac{\eta(1-p_o)(1-p_f)}{2\pi p_c j}\int_{\sigma-j\infty}^{\sigma+j\infty}\left[\frac{1-f^*(s)}{s}\right]^2 \frac{1}{1-(1-p_f)f^*(s)} \\
&\quad \times \int_0^\infty e^{-(-s+z)t_c} f_c(t_c)dt_c ds \\
&= \frac{\eta(1-p_o)(1-p_f)}{2\pi p_c j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{[1-f^*(s)]^2}{s^2[1-(1-p_f)f^*(s)]} f_c^*(-s+z)ds \\
&= -\frac{\eta(1-p_o)(1-p_f)}{p_c}\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{[1-f^*(s)]^2}{s^2[1-(1-p_f)f^*(s)]} f_c^*(-s+z). \qquad (3.34)
\end{aligned}$$

From (3.33) and (3.34), we finally obtain

Theorem 3.4. For a PCS network, the Laplace transform of the density function of the effective call holding time of a complete call is given by

$$g_c^*(z) = -\left(\frac{1-p_o}{p_c}\right)\left[\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{s-\eta(1-f^*(s))}{s^2} f_c^*(-s+z)\right.$$

$$+\eta(1-p_f)\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{[1-f^*(s)]^2 f_c^*(-s+z)}{s^2[1-(1-p_f)f^*(s)]}\Bigg]. \tag{3.35}$$

Suppose now that the call holding times are Erlang distributed with parameter (m,α) with Laplace transform $f_c(s)=(\alpha/(s+\alpha))^m$. Let

$$\begin{aligned} g_2(s) &= \frac{s-\eta(1-f^*(s))}{s^2}, & (3.36)\\ g_3(s) &= \frac{[1-f^*(s)]^2}{s^2[1-(1-p_f)f^*(s)]}. & (3.37)\end{aligned}$$

Then, we have

$$\begin{aligned} &\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{s-\eta(1-f^*(s))}{s^2}f_c^*(-s+z) = \operatorname*{Res}_{s=z+\alpha} g_2(s)\left(\frac{\alpha}{-s+z+\alpha}\right)^m \\ = \ & (-\alpha)^m \operatorname*{Res}_{s=z+\alpha}\frac{g_2(s)}{(s-(z+\alpha))^m} = \frac{(-\alpha)^m}{(m-1)!}g_2^{(m-1)}(z+\alpha),\end{aligned}$$

and

$$\operatorname*{Res}_{\substack{s=z+p\\ p\in\sigma_c}}\frac{[1-f^*(s)]^2 f_c^*(-s+z)}{s^2[1-(1-p_f)f^*(s)]} = (-\alpha)^m\operatorname*{Res}_{s=z+\alpha}\frac{g_3(s)}{(s-(z+\alpha))^m} = \frac{(-\alpha)^m}{(m-1)!}g_3^{(m-1)}(z+\alpha).$$

From these and Theorem 3.4, we obtain

Theorem 3.5. For a PCS network with Erlang distributed call holding times, the Laplace transform of the density function of the effective call holding time of a complete call is given by

$$g_c^*(z) = \frac{(-1)^{m-1}\alpha^m(1-p_o)}{(m-1)!p_c}\left[g_2^{(m-1)}(z+\alpha)+(1-p_f)\eta g_3^{(m-1)}(z+\alpha)\right], \tag{3.38}$$

the expected effective call holding time of a complete call is given by

$$T_c = -g_c^{(1)}(0) = \frac{(-1)^m\alpha^m(1-p_o)}{(m-1)!p_c}\left[g_2^{(m)}(\alpha)+(1-p_f)\eta g_3^{(m)}(\alpha)\right], \tag{3.39}$$

and the variance of the effective complete call holding times is given by

$$\begin{aligned} V_c \ &= \ \frac{(-1)^{m-1}\alpha^m(1-p_o)}{(m-1)!p_c}\left[g_2^{(m+1)}(\alpha)+(1-p_f)\eta g_3^{(m+1)}(\alpha)\right] \\ &\quad -\frac{\alpha^{2m}(1-p_o)^2}{[(m-1)!p_c]^2}\left[g_2^{(m)}(\alpha)+(1-p_f)\eta g_3^{(m)}(\alpha)\right]^2. & (3.40)\end{aligned}$$

Remark: In fact, all moments of the effective call holding time of a complete call can be obtained from (3.38) and are given as follows: (let $T_c(k)$ denote the kth moment)

$$T_c(k) = \frac{(-1)^{m+k-1}\alpha^m(1-p_o)}{(m-1)!p_c}\left[g_2^{(m+k-1)}(\alpha) + (1-p_f)\eta g_3^{(m+k-1)}(\alpha)\right], \ k \geq 1.$$

When $m = 1$, i.e., the call holding times are exponentially distributed, we have

$$\begin{aligned} T_c &= -\frac{\alpha(1-p_o)}{p_c}\left[g_2^{(1)}(\alpha) + \eta(1-p_f)g_3^{(1)}(\alpha)\right] \\ = &\quad \frac{(1-p_o)\{\alpha - \eta[\alpha f^{*(1)}(\alpha) + 2(1-f^*(\alpha))]\}}{p_c\alpha^2} \\ &+\frac{\eta(1-p_o)(1-p_f)[1-f^*(\alpha)]}{p_c\alpha[1-(1-p_f)f^*(\alpha)]}\left\{\frac{2[1-f^*(\alpha)]}{\alpha} + f^{*(1)}(\alpha) + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)}\right\} \end{aligned}$$

This is obtained in [39] using a different approach. This formula can be further simplified using (3.8) as follows:

$$\begin{aligned} T_c &= \frac{(1-p_o)}{p_c}\left\{\frac{1}{\alpha} - \frac{\eta}{\alpha}\left[f^{*(1)}(\alpha) + \frac{2(1-f^*(\alpha))}{\alpha}\right]\right\} + \frac{\eta(1-p_o)}{p_c\alpha} \\ &\quad \times \left[1 - \frac{p_f}{1-(1-p_f)f^*(\alpha)}\right]\left\{\frac{2[1-f^*(\alpha)]}{\alpha} + f^{*(1)}(\alpha) + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)}\right\} \\ &= \frac{1-p_o}{p_c}\left\{\frac{1}{\alpha} - \frac{\eta p_f}{\alpha[1-(1-p_f)f^*(\alpha)]}\left[\frac{2(1-f^*(\alpha))}{\alpha} + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)}\right]\right\} \\ &= \frac{1-p_o}{p_c\alpha}\left\{1 - \frac{\eta p_f(1-f^*(\alpha))}{\alpha[1-(1-p_f)f^*(\alpha)]}\right\} - \frac{(1-p_o)\eta p_f}{p_c\alpha} \\ &\quad \times \left[\frac{(1-f^*(\alpha))}{\alpha} + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)}\right] \\ &= \frac{1}{\alpha} - \frac{(1-p_o)\eta p_f}{p_c\alpha}\left[\frac{(1-f^*(\alpha))}{\alpha} + \frac{p_f f^{*(1)}(\alpha)}{1-(1-p_f)f^*(\alpha)}\right]. \end{aligned} \tag{3.41}$$

It is observed from simulation in [39] that $T_c < \frac{1}{\alpha}$ when the cell residence times are Gamma distributed. Using this equation, we can prove this analytically for general cell residence distribution. We want to show that the term $[\cdot]$ in the last equation of (3.41) is nonnegative. Let

$$\Delta(\alpha) = [1-f^*(\alpha)][1-(1-p_f)f^*(\alpha)] + \alpha p_f f^{*(1)}(\alpha).$$

Since $f^*(0) = 1$, $\Delta(0) = 0$. Also, we have

$$\frac{d}{d\alpha}\Delta(\alpha) = 2(1-p_f)(1-f^*(\alpha))[-f^{*(1)}(\alpha)] + \alpha p_f f^{(2)}(\alpha) > 0,$$

where we have used that $0 \le f^*(\alpha) \le 1$, $f^{*(1)}(\alpha) \le 0$ and $f^{*(2)}(\alpha) \ge 0$. Thus, $\Delta(\alpha)$ is an increasing function, hence $\Delta(\alpha) \ge \Delta(0) = 0$, from which we conclude that the term $[\cdot]$ in the last equation of (3.41) is nonnegative. Thus, for any cell residence distribution, the expected effective call holding time of a complete call is always less than the expected call holding time.

When $m > 1$, the closed-form formula may be quite involved. However, as before, recursive formulae can be easily developed for the computation of $g_2^{(p)}(\alpha)$ and $g_3^{(p)}(\alpha)$. For $g_2^{(p)}(\alpha)$, we have

$$\begin{aligned}
g_2^{(0)}(\alpha) &= \frac{\alpha - \eta(1-f^*(\alpha))}{\alpha^2} \\
g_2^{(1)}(\alpha) &= \frac{1 + \eta f^{*(1)}(\alpha) - 2\alpha g_2^{(0)}(\alpha)}{\alpha^2} \\
g_2^{(p)}(\alpha) &= \frac{\eta f^{*(p)}(\alpha) - 2p\alpha g_2^{(p-1)}(\alpha) - p(p-1)g_2^{(p-2)}(\alpha)}{\alpha^2}, \quad p \ge 2.
\end{aligned} \tag{3.42}$$

For $g_3^{(p)}(\alpha)$, let $h(s) = s^2[1-(1-p_f)f^*(s)]$ as defined in (3.11), $h^{(p)}(\alpha)$ can be computed as in (3.13), and

$$\begin{aligned}
g_3^{(0)}(\alpha) &= \frac{[1-f^*(\alpha)]^2}{h(\alpha)} \\
g_3^{(p)}(\alpha) &= \frac{-2f^{*(p)}(\alpha) + \sum_{j=0}^{p}\binom{p}{j} f^{*(j)}(\alpha) f^{*(p-j)}(\alpha) - \sum_{j=0}^{p-1}\binom{p}{j} g_3^{(j)}(\alpha) h^{(p-j)}(\alpha)}{h(\alpha)}, \\
& \quad (p \ge 1).
\end{aligned} \tag{3.43}$$

As observed earlier, we can analytically prove that for exponentially distributed call holding times the expected effective call holding times for both a complete call and an incomplete call are less than the expected call holding time. However, it is difficult to analytically prove from the formula for $m > 1$ that the expected effective call holding times for both a complete and an incomplete call are less than the expected call holding time $1/\mu$. The following intuitive argument can be used to prove this: the effective call holding times are just the "interrupted" call holding times, hence should be "smaller" than the "non-interrupted" call holding times. We will verify this observation in the next section.

3.4 Illustrations and Discussions

In this section, we present an example to show how to apply our results to analyze the effective call holding times and attempt to draw some general conclusions. In this section, we assume that the cell residence times are Gamma distributed with the following density function and Laplace transform

$$f(t) = \frac{(\gamma\eta)^\gamma t^{\gamma-1} e^{-\gamma\eta t}}{\Gamma(\gamma)}, \ t \geq 0, \ f^*(s) = \left(\frac{\gamma\eta}{s+\gamma\eta}\right)^\gamma,$$

which have mean $1/\eta$ and variance $V = 1/(\gamma\eta^2)$. The call holding times are Erlang distributed with the following density function and Laplace transform

$$f_c(t) = \frac{(m\mu)^m t^{m-1} e^{-m\mu t}}{(m-1)!}, \ t \geq 0, \ f_c^*(s) = \left(\frac{m\mu}{s+m\mu}\right)^m,$$

which have mean $1/\eta$ and variance $V = 1/(m\mu^2)$. We will investigate how the probability of a call completion and the expected effective call holding times are affected by some parameters such as the means and variances. In this section, we use the commonly used average call holding time $1/\mu = 1.76$ minutes ([34]), we use $m = 2$ when we vary γ while we use $\gamma = 1.5$ when we vary m.

As we mentioned earlier, the lognormal distributions has been proved to be a good approximation to the wireline call holding time distribution ([3]). Statistically, both the lognormal distributions and Gamma distributions (including Erlang distributions) have the same capacity to approximate the measure data ([27]), the assumption of Erlang call holding times seems to be reasonable replacement of the lognormal distributions for the performance analysis. An important advantage of the Erlang/Gamma

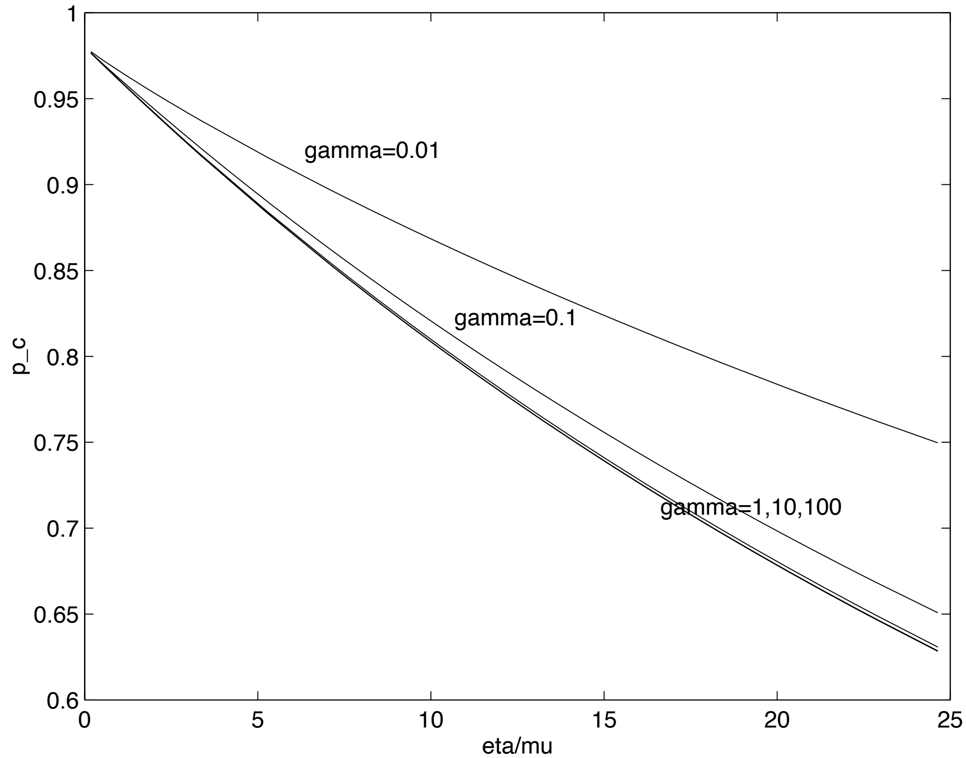


Figure 3.3: The probability of a call completion: varying γ

distribution over the lognormal distribution is that the Erlang/Gamma distribution has a simple Laplace transform, which is a desired property in our modeling.

We first consider the probability of a complete call.

Figure 3.3 shows the probability of a call completion vs. the inverse of the expected cell residence time normalized by μ, i.e., the mobility η/μ, for different values of the shaping parameter γ in the Gamma distribution, while Figure 3.4 demonstrates the variability of the call completion probability against the mobility η/μ for various values of the shaping parameter m in the Erlang distribution of the call holding times. When η is fixed, the variance of the cell residence times (the call holding times, respectively) is uniquely determined by the shaping parameter γ (m). From Figure 3.3 and Figure 3.4, we can observe the following properties:

a). The call completion probability p_c is decreasing as the expected cell residence

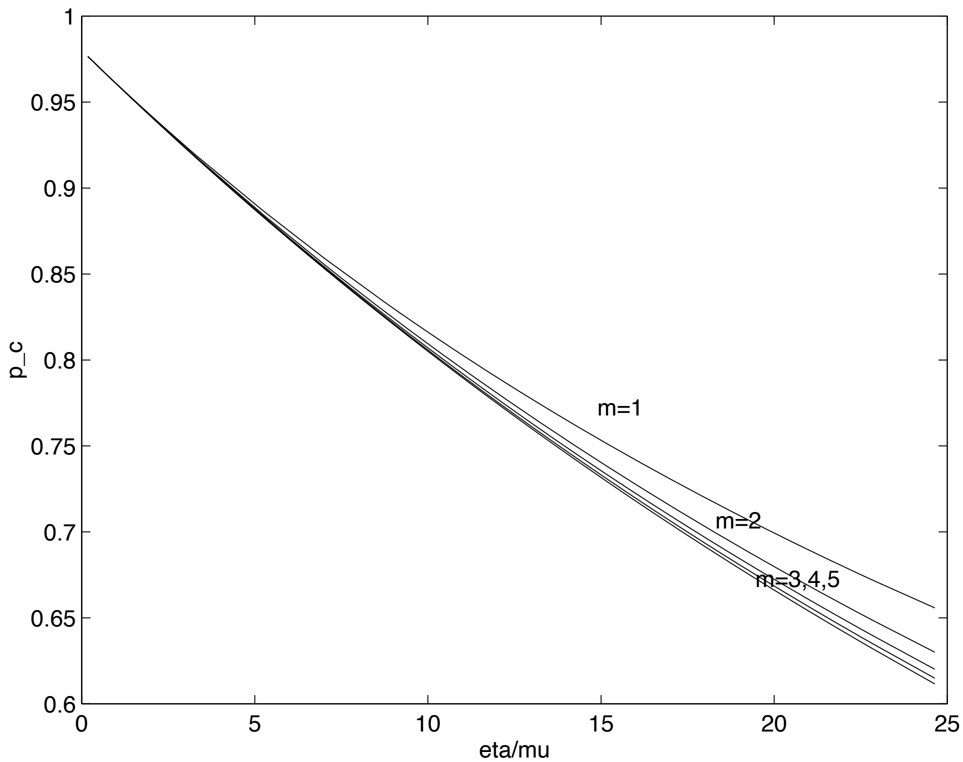


Figure 3.4: The probability of a call completion: varying m

time $(1/\eta)$ is decreasing. This is reasonable because for a fixed mobile route, i.e., the distribution of the cell residence time is fixed, then the expected cell residence time is fixed, and the longer the expected call holding time, the more often the handovers happen, the more chance the call will be dropped. This is equivalent to saying that for fixed call holding time pattern (or the fixed expected call holding time), the probability of dropping the call is increasing as the expected cell residence time become smaller, hence the probability of a call completion is decreasing as the expected cell residence time is increasing.

b). The call completion probability is decreasing as the variance of the cell residence times or the variance of the call holding times is decreasing (i.e., as γ or m is increasing).

c). When the expected cell residence time is large (i.e., when η is small), the effect of the variance of the call holding times on the call completion probability is not significant. This is why the call completion probability p_c alone is not a good evaluation measure for a PCS network, and the effective call holding times are needed.

d). There is a significant difference between the p_c for $m = 1$ and that for $m > 1$, this may be due to the fact that the exponentially distributed call holding times $(m = 1)$ is memoryless.

Next, we study the effective call holding times.

Figure 3.5 and 3.6 show the expected effective call holding time of an incomplete call. From these figures we obtain the following observations for the expected effective call holding times of an incomplete call.

(1). $T_i\mu \leq 1$, i.e., $T_i \leq 1/\mu$. This implies that the expected effective call holding time for an incomplete call is no more than the expected non-interrupted call

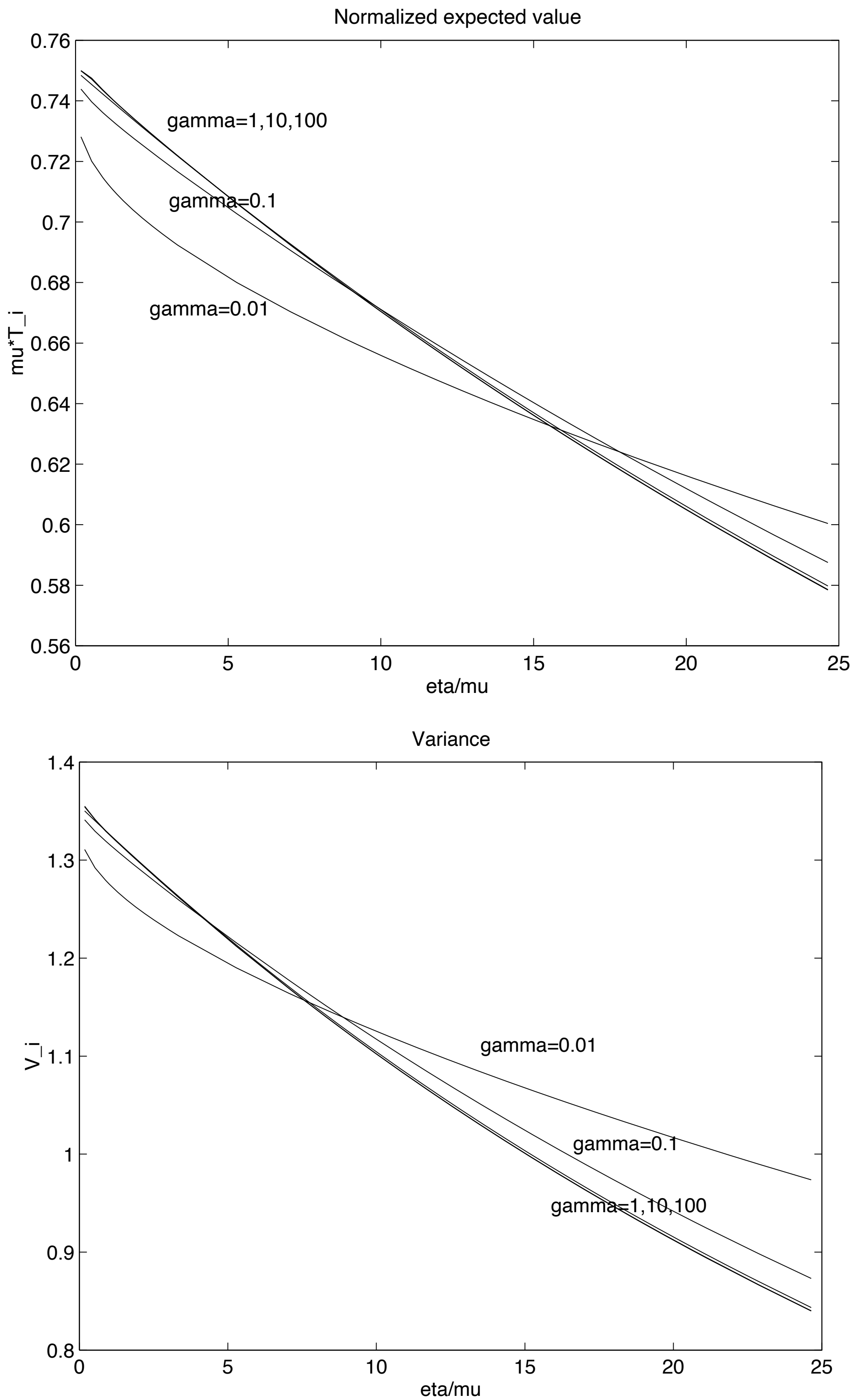


Figure 3.5: Effective call holding time of an incomplete call: varying γ

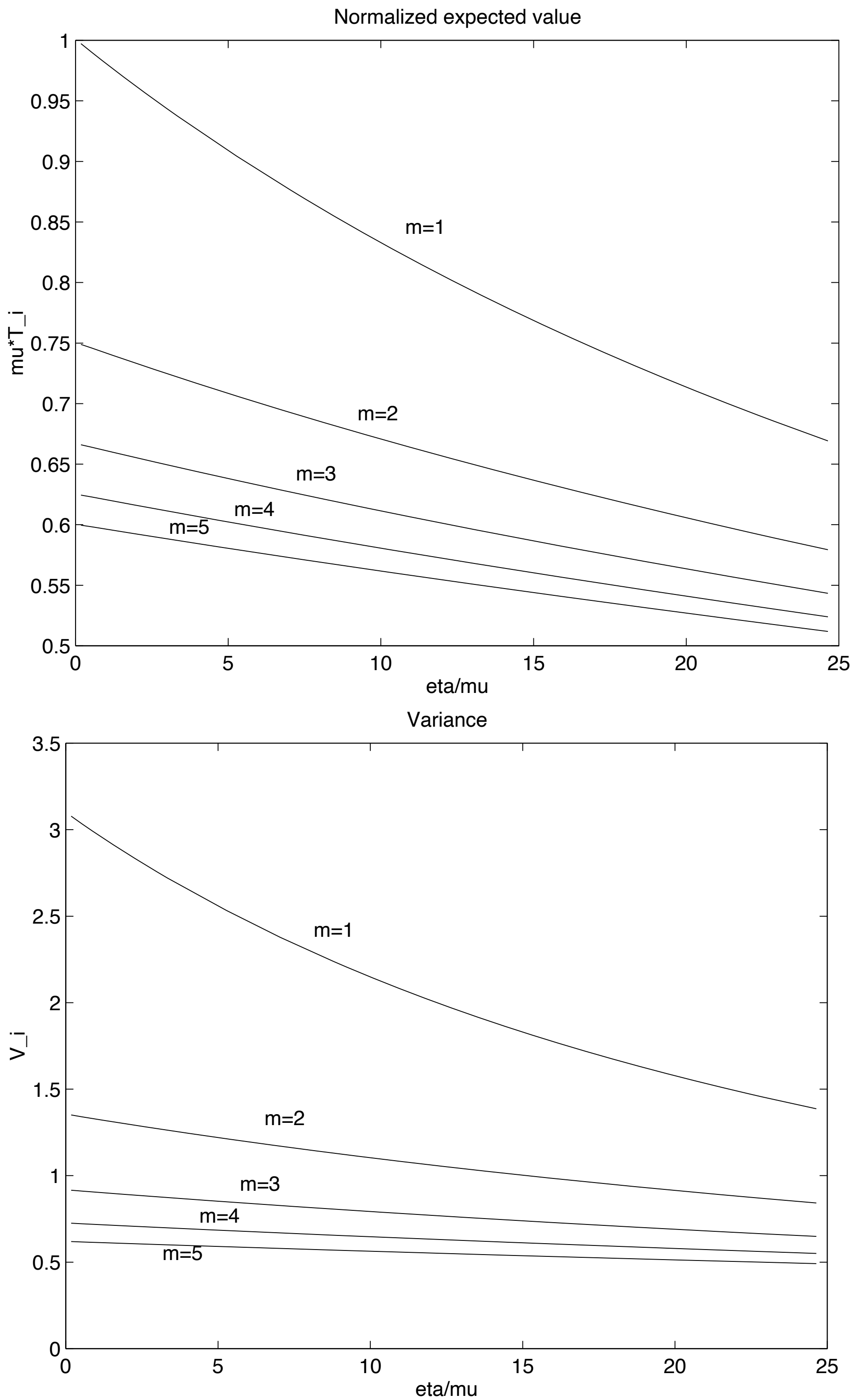


Figure 3.6: Effective call holding time of an incomplete call: varying m

holding time, which is consistent with the observation we made based on our intuition.

(2). T_i is decreasing as the mobility parameter η/μ is increasing. This is intuitive because when η/μ increases, the cell residence times decrease, more handovers are undertaken, more often the call is incomplete, hence the incomplete call holding times tend to be shorter.

(3). For small mobility η/μ, the expected incomplete call holding time is decreasing as the variance of the cell residence times is increasing; while for large mobility η/μ, this is reversed.

(4). T_i is decreasing as the variance of the call holding times is decreasing for a longer range of mobility. This relationship will be expected to reverse for very large mobility (which may be not practical range).

(5). There is a big difference between the case for $m = 1$ (the exponential distribution case) and the case for $m > 1$, which may be contributed from the memoryless property of the exponential distribution.

(6). The variance of the effective call holding time of an incomplete call is decreasing as the mobility is increasing; for small mobility η/μ, it is increasing as the variance of the cell residence times is decreasing, however, for large mobility η/μ, it is decreasing as the variance of cell residence times is decreasing. It is always decreasing as the variance of the call holding times is decreasing.

Finally, we observe the expected effective call holding time of a complete call.

Figure 3.7 and 3.8 show the results for the effective complete call holding times. We have the following observations.

(1). $T_c\mu \leq 1$, i.e., $T_c \leq 1/\mu$. This implies that the expected effective call holding time for a complete call is no more than the expected non-interrupted call holding

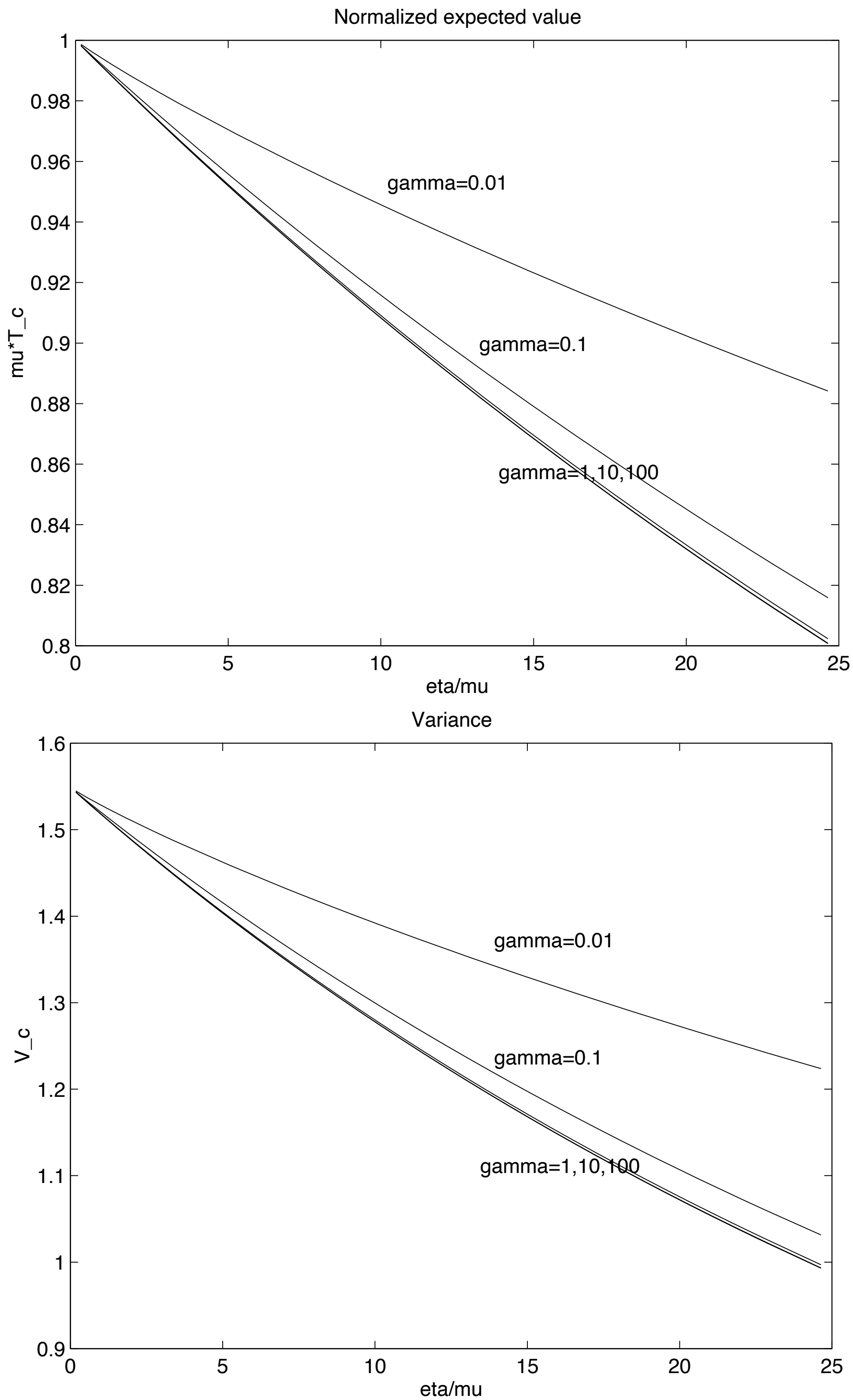


Figure 3.7: Effective call holding time of a complete call: varying γ

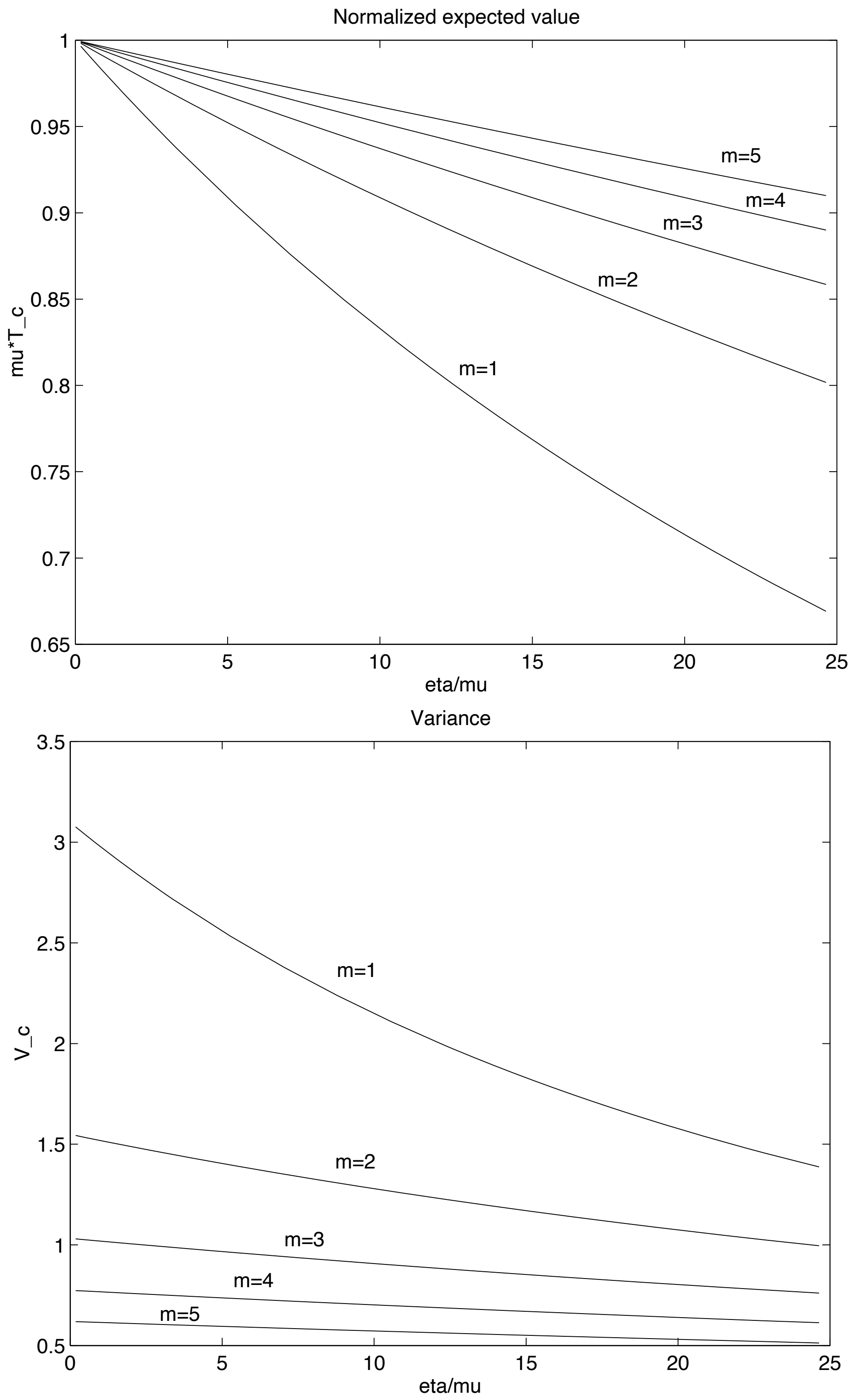


Figure 3.8: Effective call holding time of a complete call: varying m

time, which is consistent with the observation we made based on our intuition.

(2). T_c is decreasing as the mobility η/μ is increasing. This is not intuitive.

(3). The expected effective complete call holding time T_c is decreasing as the variance of the cell residence times is decreasing.

(4). T_c is increasing as the variance of the call holding times is decreasing.

(5). There is a major difference between the case for $m = 1$ (the exponential distribution case) and the case for $m > 1$, which may be contributed from the memoryless property of the exponential distribution.

(6). The variance of the effective call holding times is always decreasing as the mobility η/μ is increasing; it is always decreasing as the variance of cell residence times is decreasing; it is always decreasing as the variance of call holding times is decreasing.

As mentioned in the introduction, the study of the effective call holding times can support the provider's billing activity. Lin and Chlamtac ([39]) investigated the billing problem for the case when call holding times are exponentially distributed. Similar conclusions can be drawn for our cases here. For details of this application of our results to pricing strategies the interested reader is referred to ([39]).

3.5 Conclusions

All previous performance studies of PCS channel allocation assumed that the call holding times are exponentially distributed. While this assumption can be justified for existing cellular systems where the wireless calls are charged based on the lengths of the call holding times, future PCS systems may exercise flat rate billing programs, and therefore a more general distribution is necessary for modeling the call holding times. In this chapter, we use a general distribution to model the call holding times and derive general formulae for the probability of a call completion and the expected effective call holding times of both a complete and incomplete calls. By specifying the call holding time distribution to be Erlang, we obtain easy-to-compute recursive formulae to compute the above performance metrics. Our results can be directly applied to pricing strategies for the emerging PCS networks. In the next chapter, we will generalize our approach to more general cases where call holding times are distributed with more general distributions.

Chapter 4

MODELING PCS NETWORKS UNDER GENERAL CALL HOLDING TIME AND CELL RESIDENCE TIME DISTRIBUTIONS

4.1 Introduction

In Chapter 3, we have studied the call completion probability and effective call holding times mostly under the exponential or Erlang call holding times. As a result of the new applications in the PCS networks such as information retrieving, on-line game playing or Internet surfing, the classical assumptions on exponential or even Erlang call holding times and cell residence times may not be appropriate for modeling the new emerging integrated services in these systems. Not only does call holding time distribution vary with the new applications, also the time a customer spends in a cell (the cell residence time) will depend on the mobility of the customer, the geographic situation, and the handoff scheme used, and needs therefore be modeled as a random variable of general distribution.

It has been shown (Lemma 3.9 in [30]) that any distribution of a nonnegative random variable can be approximated by a sequence of the average summations of Erlang distributions. Therefore, it suffices to use the mixtures of Erlang distributions to model the call holding time distribution.

This chapter is the continuation of Chapter 3. Here we propose a general model which assumes that the cell residence times have general non-lattice distribution, and the call holding times are distributed with general distributions such as Gamma, hyper-exponential and hyper-Erlang distributions, as discussed later. While the advantages of such distributions are clear in that they reflect the emerging services, applying these distributions to the model is a non-trivial task. This chapter shows how to accommodate general call holding time distributions in previously proposed analytic model in Chapter 3. The following general assumptions is repeated here for convenience:

- The call arrivals form a Poisson process;
- The cell residence times are independent, identically distributed (iid) with non-lattice distribution;
- The call holding times are independent, identically distributed (iid) with non-lattice distribution.

Based on these assumptions, we obtain general formulae for the call completion probability and the distribution (its Laplace transforms) of the effective call holding times of both the complete and incomplete calls, from which expected effective call holding times can be obtained. We derive computable formulae for the cases when the call holding times are distributed according to Gamma, staged exponential or

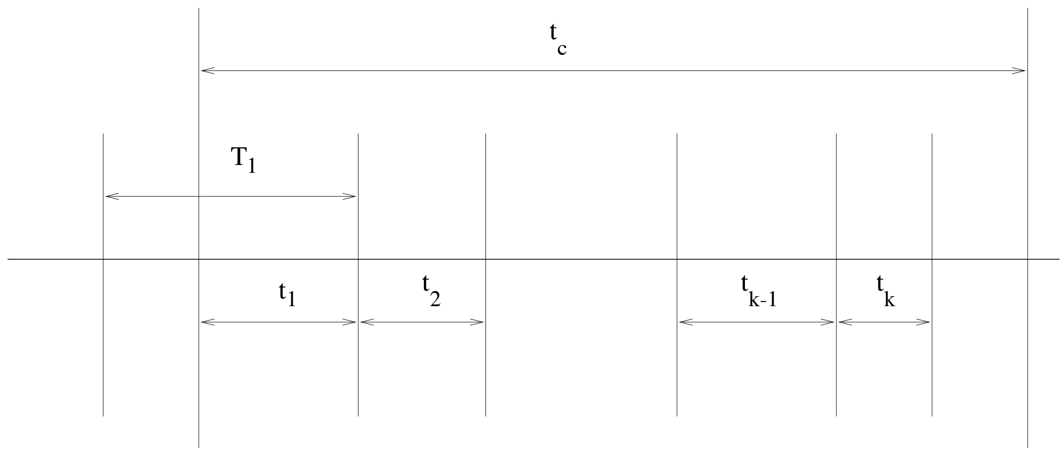


Figure 4.1: The timing diagram for a forced terminated call at kth handoff

Erlang, hyperexponential and hyper-Erlang distributions. As shown in ([39]), the analysis of the call completion probabilities and effective call holding times can provide the necessary guidelines for network performance evaluation tuning and importantly designing billing schemes in the future PCS networks ([3]). We briefly discuss a few billing rate (service charging) planning using the call completion probability and effective call holding times.

4.2 Call Completion Probability

In this section, we study the call completion probability. Our previous work ([39] and [12]) obtained formulae for the call completion probability for a PCS network with a general residence time distribution and Erlang (exponential) call holding time distribution. Here, we give further results for cases when the call holding times have other distributions.

We first consider the effective call holding time t for an incomplete call. Figure 4.1 illustrates the timing diagram for the call holding time, T_1 is the time that the portable resides at cell 1, and t_i $(i \geq 2)$ is the residence time at cell i. According to our assumptions, $T_1, t_2, \ldots, t_k, \ldots$ are iid. Let t_i have non-lattice density function

$f(\cdot)$ with the mean $1/\eta$ and $f^*(s)$ be the Laplace transform of $f(\cdot)$ (we will use $*$ to denote the Laplace transform following the tradition [32]). Suppose that a call for the portable occurs when the portable is in cell 1. Let t_1 be the interval between the time instant when the call arrives and when the portable moves out of cell 1. Let $r(t_1)$ and $r^*(s)$ be the density function and the Laplace transform of t_1 distribution, respectively. From the renewal theory ([32]), t_1 is the residual life of the cell residence time of the portable in cell 1, so we have

$$
\begin{aligned}
r(t_1) &= \eta \int_{t_1}^{\infty} f(\tau) d\tau && (4.1)\\
r^*(s) &= \frac{\eta}{s}[1 - f^*(s)] && (4.2)
\end{aligned}
$$

Let $t = t_1 + t_2 + \cdots + t_k$ be the effective holding time, and $f_k(t)$ and $f_k^*(s)$ be its density function and Laplace transform. Since $t_1, t_2, \ldots, t_k$ are independent, it is easy to derive

$$
f_k^*(s) = E[\exp(-s \sum_{i=1}^{k} t_i)] = r^*(s)[f^*(s)]^{k-1} = \frac{\eta}{s}(1 - f^*(s)) [f^*(s)]^{k-1}. \qquad (4.3)
$$

Let p_o be the probability that a new call attempt is blocked (i.e., the call is never connected), p_c be the probability that a call is completed (i.e., the call is connected and completed), p_f be the forced termination probability or the probability that no radio channel is available when a handoff call arrives. Then the probability of an incomplete call (i.e., the call is connected but is eventually forced to terminate) p_i is $1 - p_o - p_c$, which is given by

$$
\begin{aligned}
1 - p_o - p_c &= \sum_{k=1}^{\infty} \left\{ \int_0^{\infty} \int_t^{\infty} (1 - p_o) f_k(t) (1 - p_f)^{k-1} p_f f_c(t_c) dt_c dt \right\} \\
&= \sum_{k=1}^{\infty} \left\{ \int_0^{\infty} (1 - p_o) f_k(t) (1 - p_f)^{k-1} p_f \left[\int_t^{\infty} f_c(t_c) dt_c \right] dt \right\}, \qquad (4.4)
\end{aligned}
$$

where $f_c(t_c)$ is the density function of the call holding times. In last chapter, we have obtained

$$
\begin{aligned}
1 - p_o - p_c &= \frac{\eta(1 - p_o)p_f}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{1 - f^*(s)}{s[1 - (1 - p_f) f^*(s)]} \cdot \frac{f_c^*(-s) - 1}{s} ds \\
&= -\eta(1 - p_o)p_f \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{(1 - f^*(s))(1 - f_c^*(-s))}{s^2[1 - (1 - p_f) f^*(s)]} ds. \qquad (4.5)
\end{aligned}
$$

Since $f^*(s)$ has no poles in the right half complex plane, $|(1-p_f)f^*(s)| < 1$. Let σ_c denote the set of poles of $f_c^*(-s)$ in the right half complex plane (i.e., $-\sigma_c = \{-z|z \in \sigma_c\}$ is the set of poles of $f_c^*(s)$ in the left half plane). In Chapter 3, we have obtained the call completion probability for the case when the call holding times are Erlang distribution. Next, we study other cases when the call holding times have other distributions.

Assume that the call holding times are Gamma distributed with the following density function

$$f_c(t) = \frac{\alpha^\gamma t^{\gamma-1} e^{-\alpha t}}{\Gamma(\gamma)},\ \gamma > 0 \tag{4.6}$$

where γ is the shape parameter and $\alpha = \gamma\mu$ is the scale parameter. This density has the following Laplace transform

$$f_c^*(s) = \left(\frac{\alpha}{s+\alpha}\right)^\gamma. \tag{4.7}$$

Substituting (4.7) into (4.5), we obtain

$$1 - p_o - p_c = -\eta(1-p_o)p_f \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{(1-f^*(s))(1-(\alpha/(-s+\alpha))^\gamma))}{s^2[1-(1-p_f)f^*(s)]} ds. \tag{4.8}$$

In order to evaluate this integral, we need the following lemma.

Lemma 4.1. Let $\alpha > \sigma > 0$ and $R(s)$ is analytic on $C_\sigma = \{s|\mathrm{Re}(s) \geq \sigma\}$ ($\mathrm{Re}(\cdot)$ denotes the real part) with $R^{(i)}(s)s \to 0$ ($i = 0, 1, \ldots, \lfloor\gamma\rfloor$) as $s \to \infty$ in C_σ ($\lfloor x \rfloor$ denotes the integral part of x for any positive real number). Then, we have

$$\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{R(s)}{(s-\alpha)^\gamma} ds = \frac{1-e^{-j2\pi\gamma}}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \int_0^\infty \frac{R^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}} ds, \tag{4.9}$$

where $(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor) = 1$ when $0 < \gamma < 1$.

Proof: If γ is not an integer, then $s = \alpha$ is a branch point of the integrand ([35]). Let us cut the complex plane on the real axis from $s = \alpha$ right to the ∞ as shown in Figure 4.2, where $C_R = \{s = \sigma + Re^{j\theta}|0 < \theta \leq \pi/2,\ 3\pi/2 \leq \theta < 2\pi\}$, $C_\rho = \{s = \alpha + \rho e^{j\theta}|0 < \theta < 2\pi\}$, $l_1 = \{s|\mathrm{Im}(s) = 0\}$ and $l_2 = \{s|\mathrm{Im}(s) = 2\pi\}$. In the domain

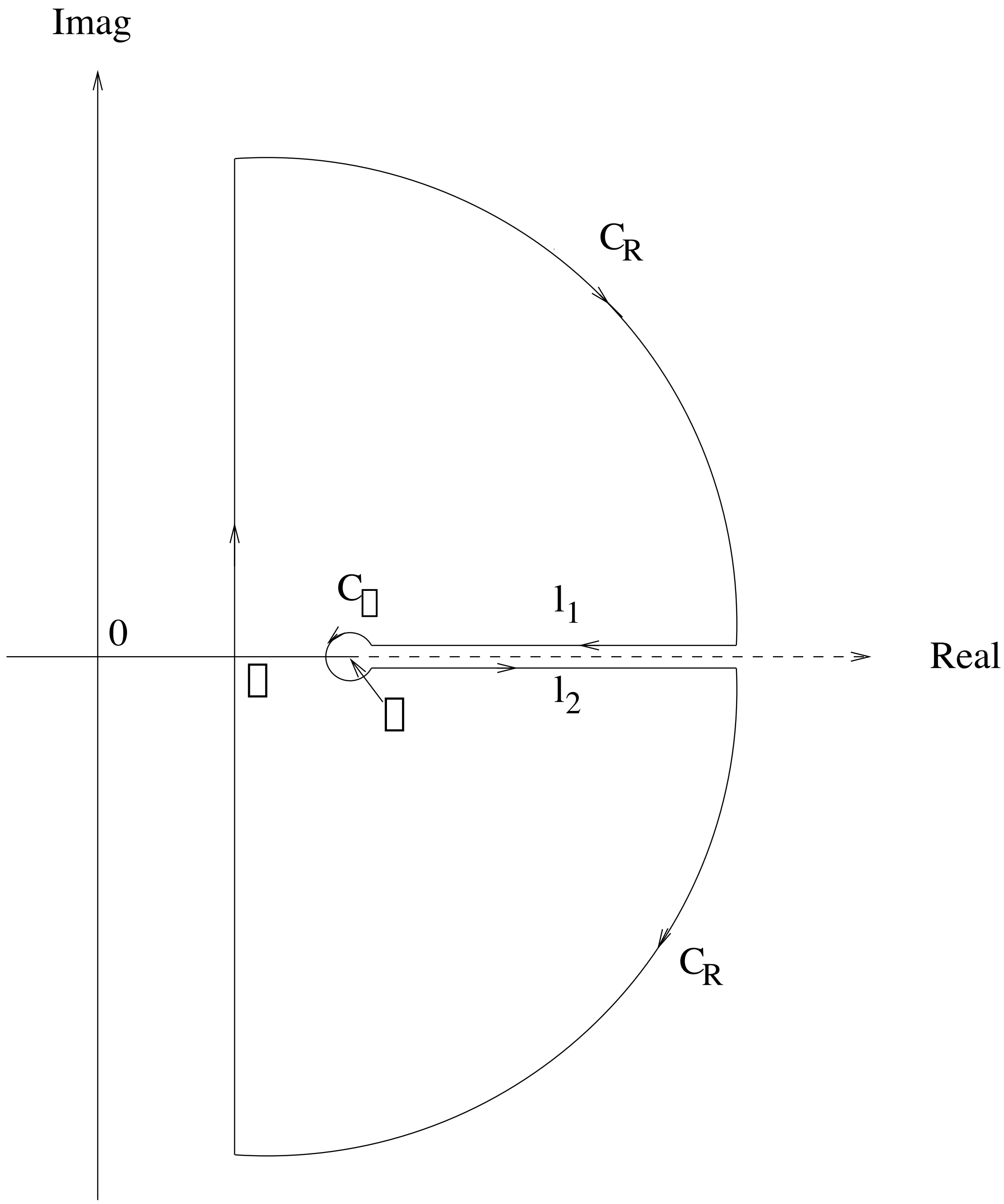


Figure 4.2: The integration contour and branching cut

enclosed by the contour $\sigma - j\infty \to \sigma + j\infty \to C_R \to l_1 \to C_\rho \to l_2 \to C_R$, the function $R(s)/(s-\alpha)^\gamma$ is analytic, hence from the Residue Theorem ([35]) we have

$$\left(\int_{\sigma-jR}^{\sigma+jR} + \int_{C_R} + \int_{l_1} + \int_{l_2} + \int_{C_\rho}\right) \frac{R(s)}{(s-\alpha)^\gamma} ds = 0,$$

from which we obtain

$$\int_{\sigma-jR}^{\sigma+jR} \frac{R(s)}{(s-\alpha)^\gamma} ds = -\left(\int_{C_R} + \int_{l_1} + \int_{l_2} + \int_{C_\rho}\right) \frac{R(s)}{(s-\alpha)^\gamma} ds. \tag{4.10}$$

Since $\lim_{s\in C_R,\ R\to\infty} R(s)s = 0$, we have

$$\lim_{R\to\infty} \int_{C_R} \frac{R(s)}{(s-\alpha)^\gamma} ds = 0.$$

On the line l_1, $(s-\alpha)^\gamma = |s-\alpha|^\gamma$, while on l_2, $(s-\alpha)^\gamma = |s-\alpha|^\gamma e^{j2\pi\gamma}$. From (4.10), by letting $R \to \infty$, we obtain

$$\begin{aligned}
&\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{R(s)}{(s-\alpha)^\gamma} ds \\
= {}& \int_{\alpha+\rho}^{\infty} \frac{R(s)}{(s-\alpha)^\gamma} ds - \frac{j}{\rho^{\gamma-1}} \int_0^{2\pi} R(a+\rho e^{j\theta}) e^{-j(\gamma-1)\theta} d\theta - e^{-j2\pi\gamma} \int_{\alpha+\rho}^{\infty} \frac{R(s)}{(s-\alpha)^\gamma} ds \\
= {}& (1-e^{-j2\pi\gamma}) \int_{\alpha+\rho}^{\infty} \frac{R(s)}{(s-\alpha)^\gamma} ds - \frac{j}{\rho^{\gamma-1}} \int_0^{2\pi} R(a+\rho e^{j\theta}) e^{-j(\gamma-1)\theta} d\theta \\
= {}& (1-e^{-j2\pi\gamma}) \int_{\rho}^{\infty} \frac{R(\alpha+s)}{s^\gamma} ds - \frac{j}{\rho^{\gamma-1}} \int_0^{2\pi} R(a+\rho e^{j\theta}) e^{-j(\gamma-1)\theta} d\theta \\
= {}& (1-e^{-j2\pi\gamma}) \left[\frac{R(\alpha+s)}{(-\gamma+1)s^{\gamma-1}}\Big|_\rho^\infty - \frac{1}{-\gamma+1} \int_\rho^\infty \frac{R^{(1)}(\alpha+s)}{s^{\gamma-1}} ds \right] - \frac{j}{\rho^{\gamma-1}} \\
& \times \left[-\frac{1}{j(\gamma-1)} R(\alpha+\rho e^{j\theta}) e^{-j(\gamma-1)\theta}\Big|_0^{2\pi} + \frac{1}{j(\gamma-1)} \int_0^{2\pi} R^{(1)}(\alpha+\rho e^{j\theta}) e^{-j(\gamma-2)\theta} d\theta \right] \\
= {}& \frac{1-e^{-j2\pi\gamma}}{(\gamma-1)} \int_\rho^\infty \frac{R^{(1)}(\alpha+s)}{s^{\gamma-1}} ds - \frac{j}{(\gamma-1)\rho^{\gamma-2}} \int_0^{2\pi} R^{(1)}(\alpha+\rho e^{j\theta}) e^{-j(\gamma-2)\theta} d\theta \\
= {}& \frac{1-e^{-j2\pi\gamma}}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor+1)} \int_\rho^\infty \frac{R^{(\lfloor\gamma\rfloor-1)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor+1}} ds \\
& - \frac{j}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor+1)\rho^{\gamma-\lfloor\gamma\rfloor}} \int_0^{2\pi} R^{(\lfloor\gamma\rfloor-1)}(\alpha+\rho e^{j\theta}) e^{-j(\gamma-\lfloor\gamma\rfloor)\theta} d\theta \\
= {}& \frac{1-e^{-j2\pi\gamma}}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \int_\rho^\infty \frac{R^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}} ds \\
& - \frac{j\rho^{1-(\gamma-\lfloor\gamma\rfloor)}}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \int_0^{2\pi} R^{(\lfloor\gamma\rfloor)}(\alpha+\rho e^{j\theta}) e^{-j(\gamma-\lfloor\gamma\rfloor-1)\theta} d\theta
\end{aligned}$$

$$= \frac{1-e^{-j2\pi\gamma}}{(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^\infty \frac{R^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds \quad (\text{letting } \rho\to 0)$$

from which the lemma is proved.

Remark: If γ is an integer, the above derivation can be reduced to the well-known result from the Residue Theorem. In this case, $s=\alpha$ is no longer a branch point but a pole with multiplicity γ, and we have

$$\begin{aligned}
&\frac{1-e^{-j2\pi\gamma}}{(\gamma-1)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^\infty \frac{R^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds \\
=\ &\frac{1}{(\gamma-1)!}\lim_{r\to\lfloor\gamma\rfloor}\frac{1-e^{-j2\pi\gamma}}{\gamma-\lfloor\gamma\rfloor}\int_0^\infty R^{(\lfloor\gamma\rfloor)}(\alpha+s)ds \\
=\ &\frac{j2\pi}{(\gamma-1)!}R^{(\gamma-1)}(\alpha+s)\Big|_0^\infty = -\frac{j2\pi}{(\gamma-1)!}R^{(\gamma-1)}(\alpha),
\end{aligned}$$

from which we can conclude that Lemma 4.1 gives the same result from the Residue Theorem given in Chapter 3.

Now, we are ready to give an expression for p_c. Let

$$g(s) = \frac{1-f^*(s)}{s^2\{1-(1-p_f)f^*(s)\}}. \tag{4.11}$$

Note that $g(s)$ is analytic on C_σ and $\lim_{s\to\infty} g(s)s = 0$ on C_σ. In fact, $g(s)$ satisfies all conditions in Lemma 4.1 and

$$\int_{\sigma-j\infty}^{\sigma+j\infty} g(s)ds = 0.$$

Taking this into consideration in (4.8) and using Lemma 4.1, we have

$$\begin{aligned}
p_c &= (1-p_o)\left\{1+\frac{\eta p_f}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{(1-f^*(s))(1-(\alpha/(-s+\alpha))^\gamma))}{s^2[1-(1-p_f)f^*(s)]}ds\right\} \\
&= (1-p_o)\left\{1-\frac{\eta p_f(-\alpha)^\gamma}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{[1-f^*(s)][1/(s-\alpha)]^\gamma}{s^2[1-(1-p_f)f^*(s)]}ds\right\} \\
&= (1-p_o)\left\{1-\frac{\eta p_f\alpha^\gamma e^{j\pi\gamma}}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{g(s)}{(s-\alpha)^\gamma}ds\right\} \\
&= (1-p_o)\left\{1-\frac{\eta p_f\alpha^\gamma e^{j\pi\gamma}(1-e^{-j2\pi\gamma})}{2\pi j(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^\infty\frac{g^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds\right\} \\
&= (1-p_o)\left\{1-\frac{\eta p_f\alpha^\gamma}{\pi(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\frac{(e^{j\pi\gamma}-e^{-j\pi\gamma})}{2j}\int_0^\infty\frac{g^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds\right\} \\
&= (1-p_o)\left\{1-\frac{\eta p_f\alpha^\gamma\sin\pi\gamma}{\pi(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^\infty\frac{g^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds\right\}.
\end{aligned}$$

Thus, we obtain

Theorem 4.1. For a PCS network with Gamma distributed calling holding times, the probability of a call completion is given by

$$p_c = (1-p_o)\left\{1 - \frac{\eta p_f \alpha^\gamma \sin \pi\gamma}{\pi(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^\infty \frac{g^{(\lfloor\gamma\rfloor)}(\alpha+s)}{s^{\gamma-\lfloor\gamma\rfloor}}ds\right\}. \qquad (4.12)$$

Remarks:

(1). When γ is a positive integer, the Gamma distribution becomes Erlang distribution. In this case,

$$\frac{\sin \pi\gamma}{\gamma-\lfloor\gamma\rfloor} \to \pi\cos\pi\gamma = \pi(-1)^\gamma \qquad (4.13)$$

From (4.13) and (4.12), we have

$$\begin{aligned} p_c &= (1-p_o)\left\{1 - \frac{(-1)^\gamma \alpha^\gamma \eta p_f}{(\gamma-1)!}\int_0^\infty g^{(\gamma)}(\alpha+s)ds\right\} \\ &= (1-p_o)\left\{1 + \frac{(-\alpha)^\gamma \eta p_f}{(\gamma-1)!} g^{(\gamma-1)}(\alpha)\right\}, \end{aligned}$$

This equation gives the same result obtained from the Residue Theorem in Chapter 3.

(2). The derivatives $g^{(p)}(\alpha+s)$ can be computed using the recursive algorithm developed in Chapter 3, and the integral in the theorem can be computed numerically either by approximation or the residue theorem if the poles of $g(s)$ (i.e., its analytic continuation) are easy to find (one special case is that when $g(s)$ is an rational function). This detail will be left to the reader.

If the call holding time has such a distribution that its Laplace transform $f_c^*(s)$ has no branch points but has possible isolated poles, then from the Residue Theorem we can obtain in Chapter 3

Theorem 4.2. If the Laplace transform of the call holding time distribution only has isolated poles in the left half of the complex plane, then the probability of a call completion is given by

$$p_c = (1-p_o)\left\{1 + \eta p_f \sum_{p\in\sigma_c} \operatorname*{Res}_{s=p} \frac{(1-f^*(s))f_c^*(-s)}{s^2\{1-(1-p_f)f^*(s)\}}\right\}, \tag{4.14}$$

where $\mathrm{Res}_{s=p}$ denotes the residue at the pole $s = p$. In particular, when $f_c^*(s)$ is rational function, then (4.14) is valid.

In last chapter, we have obtained the formula for the case that the call holding times are Erlang distributed. It is well-known ([32]) that the Erlang distribution can be obtained by a series of independent, identically distributed random variables. By serial-parallel stages, a large number of general distributions can be obtained from exponential distributed random variables ([32]). It is easy to observe that when $f_c^*(s)$ is a rational function, then Theorem 4.2 can be easily applied to find p_c. The distributions obtained by the method of stages belong to this class. We will discuss some of the important cases next.

Let the call holding times be distributed according to the $r-$stage exponential distribution (the generalized Erlang distribution) with parameters $\mu_1, \mu_2, \ldots, \mu_r$ (which are distinct) whose Laplace transform is ([32])

$$f_c^*(s) = \left(\frac{\mu_1}{s+\mu_1}\right)\left(\frac{\mu_2}{s+\mu_2}\right)\cdots\left(\frac{\mu_r}{s+\mu_r}\right). \tag{4.15}$$

A random variable with this distribution is in fact the summation of r exponentially distributed random variables with the parameters $\mu_1, \mu_2, \ldots, \mu_r$. From this, we have

Corollary 4.1. For a PCS network with $r-$stage exponential distribution of distinct positive parameters $\mu_1, \mu_2, \ldots, \mu_r$, the probability of a call completion is

$$p_c = (1-p_o)\left\{1 + (-1)^r \eta p_f \sum_{i=1}^{r} \left(\prod_{j\neq i} \frac{\mu_j}{\mu_i - \mu_j}\right) \mu_i g(\mu_i)\right\}. \tag{4.16}$$

Proof: From Theorem 4.2, we obtain

$$p_c \;=\; (1-p_o)\left\{1 + (-1)^r \eta p_f \sum_{i=1}^{r} \operatorname*{Res}_{s=\mu_i} \frac{g(s)\mu_1\cdots\mu_r}{(s-\mu_1)\cdots(s-\mu_r)}\right\}$$

$$
\begin{aligned}
&= (1-p_o)\left\{1+(-1)^r\eta p_f\sum_{i=1}^{r}\operatorname*{Res}_{s=\mu_i}\frac{g(s)\mu_1\cdots\mu_r}{(s-\mu_1)\cdots(s-\mu_r)}\right\}\\
&= (1-p_o)\left\{1+(-1)^r\eta p_f\sum_{i=1}^{r}\left(\prod_{j\neq i}\frac{\mu_j}{\mu_i-\mu_j}\right)\mu_i g(\mu_i)\right\}.
\end{aligned}
$$

This completes the proof.

Remark: Corollary 4.1 is also valid when $\mu_1 = \mu_2 = \cdots = \mu_r = \alpha$, i.e., the Erlang distribution, in which case a limit has to be taken. In fact, we only need to show from Chapter 3 that

$$
\lim_{\substack{\mu_i\to\alpha\\ i=1,\ldots,r}}\sum_{i=1}^{r}\left(\prod_{j\neq i}\frac{\mu_j}{\mu_i-\mu_j}\right)\mu_i g(\mu_i) = \frac{\alpha^r g^{(r-1)}(\alpha)}{(r-1)!}. \tag{4.17}
$$

We use the induction to prove this. For $r = 1$, the left hand side (LHS) of (4.17) is equal to the right hand side (RHS) of (4.17). Suppose that it is true for $r-1$, then for r, we have

$$
\begin{aligned}
&\text{LHS of (4.17)}\\
&= \lim_{\mu_r\to\alpha}\left\{\lim_{\substack{\mu_i\to\alpha\\ i=1,\ldots,r-1}}\left(\sum_{i=1}^{r-1}\left(\prod_{j\neq i}^{r-1}\frac{\mu_j}{\mu_i-\mu_j}\right)\mu_i\left(\frac{g(\mu_i)\mu_r}{\mu_i-\mu_r}\right)+\prod_{j=1}^{r-1}\frac{\mu_j}{\mu_r-\mu_j}\mu_r g(\mu_r)\right)\right\}\\
&= \lim_{\mu_r\to\alpha}\left\{\frac{\alpha^{r-1}}{(r-2)!}\frac{d^{r-2}}{dr^{r-2}}\left(\frac{g(x)\mu_r}{x-\mu_r}\right)\Big|_{x=\alpha}+\left(\frac{\alpha}{\mu_r-\alpha}\right)^{r-1}\mu_r g(\mu_r)\right\}\ \text{(assumption)}\\
&= \lim_{\mu_r\to\alpha}\left\{\frac{\alpha^{r-1}\mu_r}{(r-2)!}\sum_{i=0}^{r-2}\binom{r-2}{i}\left(\frac{1}{x-\mu_r}\right)^{(i)}\Big|_{x=\alpha}g^{(r-2-i)}(\alpha)+\frac{\alpha^{r-1}\mu_r}{(\mu_r-\alpha)^{r-1}}g(\mu_r)\right\}\\
&= \lim_{\mu_r\to\alpha}\left\{\frac{\alpha^r}{(r-2)!}\sum_{i=0}^{r-2}\binom{r-2}{i}(-1)^i i!\frac{g^{(r-2-i)}(\alpha)}{(\alpha-\mu_r)^{i+1}}+\frac{\alpha^r}{(\mu_r-\alpha)^{r-1}}g(\mu_r)\right\}\\
&= \alpha^r\lim_{\mu_r\to\alpha}\frac{g(\mu_r)-\sum_{i=0}^{r-2}\frac{g^{(i)}(\alpha)}{i!}g^{(i)}(\alpha)(\mu_r-\alpha)^i}{(\mu_r-\alpha)^{i-1}}\\
&= \alpha^r\frac{g^{(r-1)}(\alpha)}{(r-1)!} = \text{RHS of (4.17)}\ \ \text{(Taylor expansion)}
\end{aligned}
$$

This proves (4.17).

Assume now that the call holding times are distributed according to the $r-$stage Erlang distribution with distinct parameters

$\mu_1, \mu_2, \ldots, \mu_r$ and positive integers $m_1, m_2, \ldots, m_r$ which has the following Laplace transform

$$f_c^*(s) = \left(\frac{\mu_1}{s+\mu_1}\right)^{m_1}\left(\frac{\mu_2}{s+\mu_2}\right)^{m_2}\cdots\left(\frac{\mu_r}{s+\mu_r}\right)^{m_r}. \tag{4.18}$$

This distribution can be obtained from the sum of r independent Erlang distributed random variables with parameters (m_i, μ_i) $(i = 1, 2, \ldots, r)$.

For this case, we have

Corollary 4.2. For a PCS network with $r-$stage Erlang distributed call holding times, we have

$$p_c = (1-p_o)\left\{1 + (-1)^{\sum_{i=1}^r m_i}\eta p_f\left(\prod_{i=1}^r \mu_i^{m_i}\right)\sum_{i=1}^r \frac{d^{m_i-1}}{ds^{m_i-1}}\left(\frac{g(s)}{\prod_{j\neq i}(s-\mu_j)^{r_j}}\right)\Big|_{s=\mu_i}\right\}. \tag{4.19}$$

Equation (4.19) is specific for serial stages. For $r-$stage parallel exponential distribution (the hyperexponential distribution) with the parameters $\mu_1, \ldots, \mu_r$ and $\alpha_1, \ldots, \alpha_r$ with the following Laplace transform ([32])

$$f_c^*(s) = \sum_{i=1}^r \alpha_i \frac{\mu_i}{s+\mu_i}, \quad \alpha_i \geq 0, \ \sum_{i=1}^r \alpha_i = 1. \tag{4.20}$$

We have

Corollary 4.3. For a PCS network with hyperexponential distributed call holding times, we have

$$p_c = (1-p_o)\left\{1 - \eta p_f \sum_{i=1}^r \alpha_i \mu_i g(\mu_i)\right\}. \tag{4.21}$$

More general cases can be obtained from serial-parallel stages. One important case ([32]) is the hyper-Erlang distribution with the following density function

$$f_c(t) = \sum_{i=1}^r \alpha_i \frac{m_i\mu_i(m_i\mu_i t)^{m_i-1}}{(m_i-1)!}e^{-m_i\mu_i t}, \quad t \geq 0 \tag{4.22}$$

with the following Laplace transform

$$f_c^*(s) = \sum_{i=1}^r \alpha_i \left(\frac{m_i\mu_i}{s+m_i\mu_i}\right)^{m_i}. \tag{4.23}$$

This distribution is obtained from r parallel Erlang distributed random variables. It is shown ([30]) that the hyper-Erlang distributions can approximate any general (non-lattice) distribution.

Corollary 4.4. For a PCS network with hyper-Erlang distributed call holding times, we have

$$p_c = (1-p_o)\left\{1 + \eta p_f \sum_{i=1}^{r} (-1)^{m_i} \frac{\alpha_i (m_i\mu_i)^{m_i}}{(m_i-1)!} g^{(m_i-1)}(m_i\mu_i)\right\}. \tag{4.24}$$

One general and interesting case obtained by the method of stages is the distribution with Laplace transform

$$f_c^*(s) = \sum_{i=1}^{r} \alpha_i \prod_{j=1}^{m_i} \left(\frac{\mu_{ij}}{s+\mu_{ij}}\right).$$

A similar result for the probability of a call completion can be derived, details are left to the reader. For the method of stages, the interested reader is referred to [32]. For the computation of $g^{(p)}(\alpha)$ needed in the above, a recursive algorithm is constructed in Chapter 3.

4.3 Expected Effective Call Holding Times

In the preceding section we discussed the probability for a call to complete. To fully characterize the performance of a PCS network it is necessary to also know the expected elapsed times for the complete and the incomplete calls (their so-called effective call holding times, respectively). In Chapter 3, we have presented a few general results for the effective call holding times of complete and incomplete calls, in particular for the case when the call holding times are Erlang distributed. This section provides new results for other interesting cases.

We first consider the effective call holding time of an incomplete call. As in Chapter 3, the density function for the effective call holding time of an incomplete call that is forced to terminate is given by

$$\begin{aligned} g_i(t) &= \left(\frac{1}{p_i}\right)\left[\sum_{k=1}^{\infty} f_k(t)(1-p_o)(1-p_f)^{k-1}p_f\int_t^{\infty} f_c(t_c)dt_c\right] \\ &= \left(\frac{1-p_o}{1-p_c-p_o}\right)\left[\sum_{k=1}^{\infty} f_k(t)(1-p_f)^{k-1}p_f\int_t^{\infty} f_c(t_c)dt_c\right] \end{aligned} \tag{4.25}$$

where $p_i = 1 - p_o - p_c$ denotes the probability of a call to be incomplete and p_c is computed in the previous section. In the first equation, the term under the summation is the density that the call is forced to terminate after k handoffs.

We want to find the Laplace transform of $g_i^*(z)$ from which the expected value can be easily obtained. As in Chapter 3, from (4.19), we can obtain

$$g_i^*(z) = \frac{\eta(1-p_o)p_f}{2p_i\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} \frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}\cdot\frac{f_c^*(-s+z)-1}{s-z}ds. \tag{4.26}$$

Since

$$\lim_{s\to z}\frac{f_c^*(-s+z)-1}{s-z} = f_c^{*(1)}(0),$$

$s = z$ is a removable singular point ([35]) of the integrand of (4.26). Thus, the poles of the integrand in the right half complex plane is those of $f_c^*(-s+z)$, i.e., $\{z+p|p\in\sigma_c\}$. Let

$$h_1(s) = sg(s) = \frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}. \tag{4.27}$$

Assume that the call holding times are Gamma distributed as in (4.6) and (4.7). Taking (4.7) into (4.26), choosing σ to be greater than the real part of z and using Lemma 4.1, we have

$$\begin{aligned} g_i^*(z) &= \frac{\eta(1-p_o)p_f}{2p_i\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{1-f^*(s)}{s[1-(1-p_f)f^*(s)]}\cdot\frac{(\alpha/(-s+z+\alpha))^{\gamma}-1}{s-z}ds \\ &= \frac{\eta(1-p_o)p_f}{2p_i\pi j}\int_{\sigma-z-j\infty}^{\sigma-z+j\infty} h_1(s+z)\cdot\frac{(-\alpha/(s+\alpha))^{\gamma}-1}{s}ds \\ &= \frac{\eta(1-p_o)p_f}{2p_i\pi j}\int_{\sigma-z-j\infty}^{\sigma-z+j\infty} h_1(s+z)\cdot\frac{(-\alpha/(s+\alpha))^{\gamma}}{s}ds \\ &= \frac{\eta(1-p_o)p_f(-\alpha)^{\gamma}}{2p_i\pi j}\int_{\sigma-z-j\infty}^{\sigma-z+j\infty}\frac{h_1(s+z)/s}{(s-\alpha)^{\gamma}}ds \\ &= \frac{\eta(1-p_o)p_f\alpha^{\gamma}e^{j\pi\gamma}(1-e^{-j2\pi\gamma})}{2p_i\pi j(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}\int_0^{\infty}\frac{\frac{d^{\lfloor\gamma\rfloor}}{dx^{\lfloor\gamma\rfloor}}\left(\frac{h_1(x+z)}{x}\right)\Big|_{x=s+\alpha}}{s^{\gamma-\lfloor\gamma\rfloor}}ds. \end{aligned} \tag{4.28}$$

Using the well-known formula

$$(uv)^{(p)} = \sum_{i=0}^{p} \binom{p}{i} u^{(i)} v^{(p-i)}, \tag{4.29}$$

we obtain the following

Theorem 4.3. For a PCS network with Gamma call holding times, the Laplace transform of the density function of the effective call holding times of an incomplete call is given by

$$g_i^*(z) = \frac{(-1)^{\lfloor\gamma\rfloor}\eta(1-p_o)p_f\alpha^\gamma\lfloor\gamma\rfloor!\sin(\pi\gamma)}{\pi(1-p_o-p_c)(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \sum_{i=0}^{\lfloor\gamma\rfloor} \frac{(-1)^i}{i!} \int_0^\infty \frac{h_1^{(i)}(s+z+\alpha)}{(s+\alpha)^{\lfloor\gamma\rfloor-i+1}s^{\gamma-\lfloor\gamma\rfloor}} ds. \tag{4.30}$$

The expected effective call holding time of an incomplete call is given by

$$T_i = \frac{(-1)^{\lfloor\gamma\rfloor+1}\eta(1-p_o)p_f\alpha^\gamma\lfloor\gamma\rfloor!\sin(\pi\gamma)}{\pi(1-p_o-p_c)(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \sum_{i=0}^{\lfloor\gamma\rfloor}(-1)^i\frac{1}{i!} \int_0^\infty \frac{h_1^{(i+1)}(s+\alpha)}{(s+\alpha)^{\lfloor\gamma\rfloor-i+1}s^{\gamma-\lfloor\gamma\rfloor}} ds. \tag{4.31}$$

Here, $(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor) = 1$ when $0 < \gamma < 1$ and the limit will be taken for $\sin(\pi\gamma)/(\gamma-\lfloor\gamma\rfloor)$ ($= \pi(-1)^\gamma$ when γ is an integer).

Proof: Since $e^{j\pi\gamma}(1-e^{-j2\pi\gamma})/(2j) = \sin(\pi\gamma)$, from (4.28) we have

$$\frac{d^p}{dx^p}\left(\frac{h_1(x+z)}{x}\right) = \sum_{i=0}^{p}(-1)^i\binom{p}{i} i!\frac{1}{x^{i+1}}h_1^{(p-i)}(x+z), \quad p = \lfloor\gamma\rfloor.$$

Remark: When γ is an integer, the Gamma distribution becomes the Erlang distribution, the above result will reduce to our previous result obtained in Chapter 3. In fact, from (4.28), when γ is an integer, we obtain

$$\begin{aligned}
g_i^*(z) &= \frac{\eta(1-p_o)p_f\alpha^\gamma\cos(\pi\gamma)}{(1-p_o-p_c)(\gamma-1)!}\int_0^\infty \frac{d^\gamma}{ds^\gamma}\left(\frac{h_1(s+z+\alpha)}{s+\alpha}\right) ds \\
&= \frac{\eta(1-p_o)p_f\alpha^\gamma\cos(\pi\gamma)}{(1-p_o-p_c)(\gamma-1)!}\frac{d^{\gamma-1}}{ds^{\gamma-1}}\left(\frac{h_1(s+z+\alpha)}{s+\alpha}\right)\Big|_0^\infty \\
&= (-1)^{\gamma-1}\frac{\eta(1-p_o)p_f\alpha^\gamma}{(1-p_o-p_c)(\gamma-1)!}\frac{d^{\gamma-1}}{ds^{\gamma-1}}\left(\frac{h_1(s+z+\alpha)}{s+\alpha}\right)\Big|_{s=0} \\
&= \frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=0}^{\gamma-1}\frac{(-\alpha)^i}{i!}h_1^{(i)}(z+\alpha).
\end{aligned}$$

If $f_c^*(-s)$ does not have any branch points in the right half complex plane, then the Residue Theorem can be used to obtain certain computable results. In fact, from (4.26) we have the following general result (see Chapter 3).

Theorem 4.4. If $f_c^*(s)$ only has isolated poles in the left half of the complex plane, then

$$g_i^*(z) = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{p\in\sigma_c} \operatorname*{Res}_{s=z+p} \frac{(1-f^*(s))f_c^*(-s+z)}{s(s-z)[1-(1-p_f)f^*(s)]}. \tag{4.32}$$

In particular, when $f_c^*(s)$ is rational function, then (4.32) is valid.

If the call holding times are $r-$stage exponentially distributed, we have the following result.

Corollary 4.5. For a PCS network with call holding times $r-$stage exponentially distributed with Laplace transform as in (4.15), then

$$g_i^*(z) = \frac{(-1)^{r+1}\left(\prod_{i=1}^r \mu_i\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \frac{h_1(z+\mu_i)}{\mu_i \prod_{j\neq i}(\mu_i-\mu_j)}, \tag{4.33}$$

and the expected effective call holding time of an incomplete call is given by

$$T_i = \frac{(-1)^{r}\left(\prod_{i=1}^r \mu_i\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \frac{h_1^{(1)}(\mu_i)}{\mu_i \prod_{j\neq i}(\mu_i-\mu_j)}. \tag{4.34}$$

Proof: By taking (4.15) into (4.32), we obtain

$$\begin{aligned} g_i^*(z) &= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{p\in\sigma_c} \operatorname*{Res}_{s=z+p} \frac{h_1(s)}{s-z} \prod_{i=1}^r \left(\frac{\mu_i}{-s+z+\mu_i}\right) \\ &= \frac{(-1)^{r+1}\left(\prod_{i=1}^r \mu_i\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \operatorname*{Res}_{s=z+\mu_i} \frac{h_1(s)/(s-z)}{\prod_{i=1}^r (s-(z+\mu_i))} \\ &= \frac{(-1)^{r+1}\left(\prod_{i=1}^r \mu_i\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \frac{h_1(z+\mu_i)}{\mu_i \prod_{j\neq i}(\mu_i-\mu_j)}, \end{aligned}$$

this proves (4.33). (4.34) can be obtained by $T_i = -g_i^{*(1)}(0)$. This completes the proof.

For $r-$ staged Erlang distributed call holding times, we have the following result.

Corollary 4.6. For a PCS network with call holding times distributed according to

$r-$stage Erlang distribution as in (4.18), we have

$$g_i^*(z) = \frac{\left(\prod_{i=1}^m \mu_i^{m_i}\right)\eta(1-p_o)p_f}{(-1)^{1+\sum_{i=1}^r m_i}(1-p_o-p_c)} \sum_{i=1}^r \frac{1}{(m_i-1)!} \frac{d^{m_i-1}}{ds^{m_i-1}} \left(\frac{h_1(s+z)}{s\prod_{j\neq i}(s-\mu_i)^{m_i}} \right)\Bigg|_{s=\mu_i}, \tag{4.35}$$

and the expected effective call holding time of an incomplete call is given by

$$T_i = \frac{\left(\prod_{i=1}^m \mu_i^{m_i}\right)\eta(1-p_o)p_f}{(-1)^{1+\sum_{i=1}^r m_i}(1-p_o-p_c)} \sum_{i=1}^r \frac{1}{(m_i-1)!} \frac{d^{m_i-1}}{ds^{m_i-1}} \left(\frac{h_1^{(1)}(s)}{s\prod_{j\neq i}(s-\mu_i)^{m_i}} \right)\Bigg|_{s=\mu_i}. \tag{4.36}$$

Proof: Taking (4.18) into (4.32), we have

$$\begin{aligned}
g_i^*(z) &= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{p\in\sigma_c} \operatorname*{Res}_{s=z+p} \frac{h_1(s)}{s-z} \prod_{i=1}^r \left(\frac{\mu_i}{-s+z+\mu_i}\right)^{m_i} \\
&= -\frac{(-1)^{\sum_{i=1}^r m_i}\left(\prod_{i=1}^m \mu_i^{m_i}\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \operatorname*{Res}_{s=z+\mu_i} \frac{h_1(s)/(s-z)}{\prod_{i=1}^r (s-(z+\mu_i))^{m_i}} \\
&= -\frac{(-1)^{\sum_{i=1}^r m_i}\left(\prod_{i=1}^m \mu_i^{m_i}\right)\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \frac{1}{(m_i-1)!} \\
&\quad \times \frac{d^{m_i-1}}{ds^{m_i-1}} \left(\frac{h_1(s)/(s-z)}{\prod_{j\neq i}(s-(z+\mu_i))^{m_i}} \right)\Bigg|_{s=z+\mu_i} \\
&= \frac{\left(\prod_{i=1}^m \mu_i^{m_i}\right)\eta(1-p_o)p_f}{(-1)^{1+\sum_{i=1}^r m_i}(1-p_o-p_c)} \sum_{i=1}^r \frac{1}{(m_i-1)!} \frac{d^{m_i-1}}{ds^{m_i-1}} \left(\frac{h_1(s+z)}{s\prod_{j\neq i}(s-\mu_i)^{m_i}} \right)\Bigg|_{s=\mu_i}.
\end{aligned}$$

This proves (4.35). (4.36) can be proved by $T_i = -g_i^{*(1)}(0)$.

For hyperexponentially distributed call holding times, we have

Corollary 4.7. For a PCS network with hyperexponentially distributed call holding times (as in (4.20), we have

$$g_i^*(z) = \frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \alpha_i h_1(z+\mu_i), \tag{4.37}$$

and the expected effective call holding time of an incomplete call is given by

$$T_i = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{i=1}^r \alpha_i h_1^{(1)}(\mu_i). \tag{4.38}$$

Proof: Taking (4.20) into (4.32), we obtain

$$g_i^*(z) = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c} \sum_{p\in\sigma_c} \operatorname*{Res}_{s=z+p} \frac{h_1(s)}{s-z} \sum_{i=1}^r \alpha_i \left(\frac{\mu_i}{-s+z+\mu_i}\right)$$

$$
\begin{aligned}
&= \frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\operatorname*{Res}_{s=z+\mu_i}\frac{h_1(s)}{s-z}\frac{\alpha_i\mu_i}{s-(z+\mu_i)} \\
&= \frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\alpha_i h_1(z+\mu_i),
\end{aligned}
$$

from which the proof can be easily completed.

For the hyper-Erlang distributed call holding times, we have

Corollary 4.8. For a PCS network with hyper-Erlang call holding times, we have

$$
g_i^*(z) = \frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\alpha_i\sum_{j=0}^{m_i-1}\frac{(-m_i\mu_i)^j}{j!}h_1^{(j)}(z+m_i\mu_i), \tag{4.39}
$$

and the expected effective call holding time of an incomplete call is given by

$$
T_i = -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\alpha_i\sum_{j=0}^{m_i-1}\frac{(-m_i\mu_i)^j}{j!}h_1^{(j+1)}(m_i\mu_i), \tag{4.40}
$$

Proof: Taking (4.23) into (4.32), we have

$$
\begin{aligned}
g_i^*(z) &= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{p\in\sigma_c}\operatorname*{Res}_{s=z+p}\left(\frac{h_1(s)}{s-z}\right)\sum_{i=1}^{r}\alpha_i\left(\frac{m_i\mu_i}{-s+z+m_i\mu_i}\right)^{m_i} \\
&= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\operatorname*{Res}_{s=z+m_i\mu_i}\left(\frac{h_1(s)}{s-z}\right)\sum_{i=1}^{r}\frac{(-1)^{m_i}\alpha_i(m_i\mu_i)^{m_i}}{(s-(z+m_i\mu_i))^{m_i}} \\
&= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\operatorname*{Res}_{s=z+m_i\mu_i}\left(\frac{h_1(s)}{s-z}\right)\frac{(-1)^{m_i}\alpha_i(m_i\mu_i)^{m_i}}{(s-(z+m_i\mu_i))^{m_i}} \\
&= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\frac{(-1)^{m_i}\alpha_i(m_i\mu_i)^{m_i}}{(m_i-1)!}\frac{d^{m_i-1}}{ds^{m_i-1}}\left(\frac{h_1(s)}{s-z}\right)\Bigg|_{s=z+m_i\mu_i} \\
&= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\frac{(-1)^{m_i}\alpha_i(m_i\mu_i)^{m_i}}{(m_i-1)!}\frac{d^{m_i-1}}{ds^{m_i-1}}\left(\frac{h_1(s+z)}{s}\right)\Bigg|_{s=m_i\mu_i} \\
&= -\frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\frac{(-1)^{m_i}\alpha_i(m_i\mu_i)^{m_i}}{(m_i-1)!} \\
&\qquad \times\sum_{j=0}^{m_i-1}\binom{m_i-1}{j}(s^{-1})^{(j)}\Big|_{s=m_i\mu_i}h_1^{(m_i-1-j)}(z+m_i\mu_i) \\
&= \frac{\eta(1-p_o)p_f}{1-p_o-p_c}\sum_{i=1}^{r}\alpha_i\sum_{j=0}^{m_i-1}\frac{(-m_i\mu_i)^j}{j!}h_1^{(j)}(z+m_i\mu_i),
\end{aligned}
$$

from which the corollary can be proved.

Next, we study the expected effective holding time for a complete call. The timing

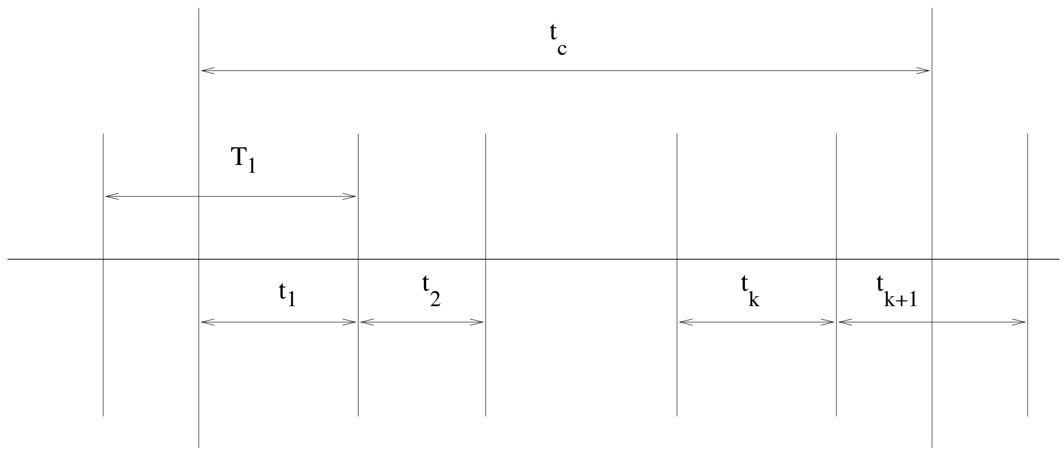


Figure 4.3: The timing diagram for the effective call times of a complete call: the call completes after k handoffs

diagram is shown in Figure 4.3, in which the call is completed when the portable is in cell k'. As before, t_c represents the effective call holding time for a complete call. If $k' = 1$, $0 \le t_c \le t_1$; while if $k' > 1$, $t_1 + t_2 + \cdots + t_{k'-1} \le t_c \le t_1 + t_2 + \cdots + t_{k'}$. Let $k = k' - 1$, then we have

$$\text{For } k = 0,\ \ 0 \le t_c \le t_1 \tag{4.41}$$

$$\text{For } k > 0,\ \ t_1 + t_2 + \cdots + t_k \le t_c \le t_1 + t_2 + \cdots + t_{k+1} \tag{4.42}$$

Using a simple conditional probability argument, we can obtain the density function $g_c(t_c)$ of the effective call holding time of a complete call is given by

$$g_c(t_c) = U(t_c) + W(t_c) \tag{4.43}$$

where

$$U(t_c) = \left(\frac{1-p_o}{p_c}\right)\left[f_c(t_c)\int_{t_c}^{\infty} r(t_1)dt_1\right], \tag{4.44}$$

$$W(t_c) = \left(\frac{1-p_o}{p_c}\right)\left[\sum_{k=1}^{\infty} f_c(t_c)\int_0^{t_c}\int_{t_c-t}^{\infty} f_k(t)(1-p_f)^k f(\tau)d\tau dt\right]. \tag{4.45}$$

$U(t_c)$ corresponds to (4.41) and $W(t_c)$ corresponds to (4.42), where $(1-p_o)$ is the probability of nonblocking, $(1-p_f)$ is the probability of no forced termination. The equation (4.44) can be derived from $P(t_c \leq x) = \sum_{k=0}^{\infty} P(t_c \leq x, k)$ where $P(t_c \leq x, k)$ denotes the probability that the call is completed in cell $k+1$ and the effective call holding time is not exceeding x. Rigorous derivation can be obtained following a similar argument in [40].

The Laplace transforms $U^*(z)$ and $W^*(z)$ of $U(t_c)$ and $W(t_c)$, respectively, are given below (see Chapter 3).

$$
\begin{aligned}
U^*(z) &= \left(\frac{1-p_o}{2\pi p_c j}\right) \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{s-\eta(1-f^*(s))}{s^2} f_c^*(z-s)ds, && (4.46) \\
W^*(z) &= \frac{\eta(1-p_o)(1-p_f)}{2\pi p_c j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{[1-f^*(s)]^2}{s^2[1-(1-p_f)f^*(s)]} f_c^*(-s+z)ds. && (4.47)
\end{aligned}
$$

Let

$$
\begin{aligned}
h_2(s) &= \frac{s-\eta(1-f^*(s))}{s^2}, && (4.48) \\
h_3(s) &= \frac{[1-f^*(s)]^2}{s^2[1-(1-p_f)f^*(s)]} = \frac{1-f^*(s)}{s^2} g(s). && (4.49)
\end{aligned}
$$

Assume that the call holding times are Gamma distributed with the Laplace transform (4.7). Then, from Lemma 4.1 we have the Laplace transform of the density function of the effective call holding time of a complete call

$$
\begin{aligned}
g_c^*(z) &= U^*(z) + W^*(z) \\
&= \left(\frac{1-p_o}{2\pi p_c j}\right) \int_{\sigma-j\infty}^{\sigma+j\infty} h_2(s) f_c^*(z-s)ds \\
&\quad + \left(\frac{\eta(1-p_o)(1-p_f)}{2\pi p_c j}\right) \int_{\sigma-j\infty}^{\sigma+j\infty} h_3(s) f_c^*(-s+z)ds \\
&= \left(\frac{1-p_o}{2\pi p_c j}\right) \int_{\sigma-j\infty}^{\sigma+j\infty} (h_2(s) + \eta(1-p_f)h_3(s)) f_c^*(-s+z)ds \\
&= \frac{1-p_o}{2\pi p_c j} \int_{\sigma-j\infty}^{\sigma+j\infty} (h_2(s) + \eta(1-p_f)h_3(s)) \left(\frac{\alpha}{-s+z}\right)^{\gamma} ds \\
&= \frac{(1-p_o)(-\alpha)^{\gamma}}{2\pi p_c j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{h_2(s) + \eta(1-p_f)h_3(s)}{(s-z-\alpha)^{\gamma}} ds \\
&= \frac{(1-p_o)\alpha^{\gamma} e^{j\pi\gamma}}{2\pi p_c j} \int_{\sigma-z-j\infty}^{\sigma-z+j\infty} \frac{h_2(s+z) + \eta(1-p_f)h_3(s+z)}{(s-\alpha)^{\gamma}} ds \\
&= \frac{(1-p_o)\alpha^{\gamma} e^{j\pi\gamma}(1-e^{-j2\pi\gamma})}{2\pi p_c j(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)}
\end{aligned}
$$

$$
\begin{aligned}
& \times \int_0^\infty \frac{h_2^{(\lfloor\gamma\rfloor)}(s+z+\alpha)+\eta(1-p_f)h_3^{(\lfloor\gamma\rfloor)}(s+z+\alpha)}{s^{\gamma-\lfloor\gamma\rfloor}}ds \\
= & \frac{(1-p_o)\alpha^\gamma\sin(\pi\gamma)}{\pi p_c(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \\
& \times \int_0^\infty \frac{h_2^{(\lfloor\gamma\rfloor)}(s+z+\alpha)+\eta(1-p_f)h_3^{(\lfloor\gamma\rfloor)}(s+z+\alpha)}{s^{\gamma-\lfloor\gamma\rfloor}}ds,
\end{aligned}
$$

from which we obtain

Theorem 4.5. For a PCS network with Gamma distributed call holding times, we have

$$
\begin{aligned}
g_c^*(z) &= \frac{(1-p_o)\alpha^\gamma\sin(\pi\gamma)}{\pi p_c(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \\
& \quad \times \int_0^\infty \frac{h_2^{(\lfloor\gamma\rfloor)}(s+z+\alpha)+\eta(1-p_f)h_3^{(\lfloor\gamma\rfloor)}(s+z+\alpha)}{s^{\gamma-\lfloor\gamma\rfloor}}ds, \qquad (4.50)
\end{aligned}
$$

and the expected effective call holding time of a complete call is given by

$$
\begin{aligned}
T_c &= -\frac{(1-p_o)\alpha^\gamma\sin(\pi\gamma)}{\pi p_c(\gamma-1)(\gamma-2)\cdots(\gamma-\lfloor\gamma\rfloor)} \\
& \quad \times \int_0^\infty \frac{h_2^{(\lfloor\gamma\rfloor+1)}(s+\alpha)+\eta(1-p_f)h_3^{(\lfloor\gamma\rfloor+1)}(s+\alpha)}{s^{\gamma-\lfloor\gamma\rfloor}}ds. \qquad (4.51)
\end{aligned}
$$

Remark: When γ is a positive integer, then Theorem 4.5 reduces to the result presented in Chapter 3 for the case when the call holding times are Erlang distributed.

When $f_c^*(s)$ does not have any branch points and has only finite isolated poles, an application of the Residue Theorem to (4.46) and (4.47) leads to the following result obtained in Chapter 3.

Theorem 4.6. For a PCS network, the Laplace transform of the density function of the effective call holding time of a complete call is given by

$$
g_c^*(z) = -\frac{1-p_o}{p_c}\left\{\sum_{p\in\sigma_c}\operatorname*{Res}_{s=z+p} h_2(s)f_c^*(-s+z) + \eta(1-p_f)\sum_{p\in\sigma_c}\operatorname*{Res}_{s=z+p} h_3(s)f_c^*(-s+z)\right\} \qquad (4.52)
$$

If the call holding times are $r-$stage exponentially distributed, then we have

Corollary 4.8. For a PCS network with the $r-$stage exponential call holding times (see equation (4.15)), the Laplace transform of the effective call holding time of a complete call is given by

$$g_c^*(z) = \frac{(-1)^{r+1}\left(\prod_{i=1}^r \mu_i\right)(1-p_o)}{p_c}\sum_{i=1}^r \frac{h_2(z+\mu_i)+\eta(1-p_f)h_3(z+\mu_i)}{\prod_{j\neq i}(\mu_i-\mu_j)}, \tag{4.53}$$

and the expected effective call holding time of a complete call is given by

$$T_c = \frac{(-1)^{r}\left(\prod_{i=1}^r \mu_i\right)(1-p_o)}{p_c}\sum_{i=1}^r \frac{h_2^{(1)}(\mu_i)+\eta(1-p_f)h_3^{(1)}(\mu_i)}{\prod_{j\neq i}(\mu_i-\mu_j)}, \tag{4.54}$$

Proof: This and the following few results can be proved as those for the effective call holding times of an incomplete call and are left to the readers.

If the call holding times are $r-$stage Erlang distributed, then we obtain the following result.

Corollary 4.9. For a PCS network with call holding times distributed according to the $r-$stage Erlang distribution (4.18), then

$$g_c^*(z) = -\frac{(1-p_o)}{p_c(-1)^{\sum_{i=1}^r m_i}}\sum_{i=1}^r \frac{\prod_{i=1}^r \mu_i^{m_i}}{(m_i-1)!}\frac{d^{m_i-1}}{ds^{m_i-1}}\left(\frac{h_2(s+z)+\eta(1-p_f)h_3(s+z)}{\prod_{j\neq i}(s-\mu_j)}\right)\Bigg|_{s=\mu_i}, \tag{4.55}$$

and the expected effective call holding time of a complete call is given by

$$T_c = \frac{(1-p_o)}{p_c(-1)^{\sum_{i=1}^r m_i}}\sum_{i=1}^r \frac{\prod_{i=1}^r \mu_i^{m_i}}{(m_i-1)!}\frac{d^{m_i-1}}{ds^{m_i-1}}\left(\frac{h_2^{(1)}(s)+\eta(1-p_f)h_3^{(1)}(s)}{\prod_{j\neq i}(s-\mu_j)}\right)\Bigg|_{s=\mu_i}. \tag{4.56}$$

If the call holding times are hyperexponentially distributed, then the following result can be obtained.

Corollary 4.10. For a PCS network with call holding times distributed according to hyperexponential distribution, then

$$g_c^*(z) = \frac{1-p_o}{p_c}\sum_{i=1}^r \alpha_i\mu_i\{h_2(z+\mu_i)+\eta(1-p_f)h_3(z+\mu_i)\}, \tag{4.57}$$

and the expected effective call holding time of a complete call is given by

$$T_c = -\frac{1-p_o}{p_c}\sum_{i=1}^r \alpha_i\mu_i\{h_2^{(1)}(\mu_i)+\eta(1-p_f)h_3^{(1)}(\mu_i)\}. \tag{4.58}$$

Finally, for the call holding times are hyper-Erlang distributed, we have

Corollary 4.11. For a PCS network with call holding times distributed according to hyper-Erlang distribution as in (4.22), we have

$$g_c^*(z) = -\frac{1-p_o}{p_c}\sum_{i=1}^{r}\alpha_i\frac{(-m_i\mu_i)^{m_i}}{(m_i-1)!}\{h_2^{(m_i-1)}(z+m_i\mu_i)+\eta(1-p_f)h_3^{(m_i-1)}(z+m_i\mu_i)\}, \tag{4.59}$$

and the expected effective call holding time of a complete call is given by

$$T_c = \frac{1-p_o}{p_c}\sum_{i=1}^{r}\alpha_i\frac{(-m_i\mu_i)^{m_i}}{(m_i-1)!}\{h_2^{(m_i)}(m_i\mu_i)+\eta(1-p_f)h_3^{(m_i)}(m_i\mu_i)\}. \tag{4.60}$$

Remarks:

(1). Although we only gave the results on the expected call holding times of a complete or an incomplete call, it is apparent to obtain their variance and higher moments from their Laplace transforms.

(2). Most results involve the computations of certain functions like $g(s)$, $h_1(s)$, $h_2(s)$ and $h_3(s)$ at certain points. The recursive algorithms developed in Chapter 3 can be used to facilitate the computations.

4.4 Service Charging Planning

In Chapter 3 and this chapter, we have found analytical expressions for the expected call holding times for both a complete call and an incomplete call. From these two quantities, we can easily compute the *expected call holding time*, i.e., the expected time of service usage for a call (either complete or incomplete). In this section, we want to show the usefulness of the effective call holding time for a complete call and the effective call holding time for an incomplete call. I show that these two quantities are not only of interest in their own rights in performance evaluation, but also very useful in the service charging planning.

PCS networks are targeted to provide integrated services radio links, the usage of the networks may be charged differently. It is not fair to charge the users for the incomplete calls the same rate as for complete calls, it is also not fair for service providers to charge the same rate for long incomplete calls as for short incomplete calls. In the last, because there are no analytical solutions for the effective call holding times for a complete call and an incomplete call, in the cellular rating systems there are no differentiation between a complete call and an incomplete call, and the same rate is used for the air time, which can be determined by the effective call holding time. Now, we can differentiate the air times for complete calls and incomplete calls, different rating systems can be developed. In the following, we discuss a few possible service charging (rating) systems for PCS network services.

4.4.1 Flat Rate Planning

The simplest rating planning will be the flat rate planning, which applies a flat rate to all calls (either complete or incomplete). This planning is easy to implement and easy to advertise, the determining factor for the charging rate is the expected effective call holding time

$$T = \frac{p_i}{1-p_o}T_i + \frac{p_c}{1-p_o}T_c = \tilde{p}_i T_i + \tilde{p}_c T_c,$$

where $\tilde{p}_i = p_i/(1-p_o)$ and $\tilde{p}_c = p_c/(1-p_o)$. Noticing that $p_o + p_i + p_c = 1$, we have $\tilde{p}_i + \tilde{p}_c = 1$. In wireline networks, there is no forced termination, $p_i = 0$, the flat rate planning is reasonable. Obviously, when the call dropping probability is sufficiently small (this is the case when either call traffic is light and mobility is comparably low or the number of channels are more than enough to support all incoming calls), then the flat rate planning is useful.

4.4.2 Partial Flat Rate Planning

The second planning is to use the partial flat rate. Due to the increase of the number of users and the mobility of users, the cell traffic will tend to increase, hence the effective call holding times for incomplete calls may be significant. It will be reasonable to use different rates for complete calls and incomplete calls. Let C_i and C_c denote the rates for an incomplete call and a complete call, respectively, then the average cost per call is given

$$C = \tilde{p}_i C_i T_i + \tilde{p}_c C_c T_c. \tag{4.61}$$

Of course, $C_i \leq C_c$. From previous section, we know that the expected effective call holding time T_i for an incomplete call is comparably longer than the expected effective call holding time T_c for a complete call. Hence, from (4.61), we observe that increasing the rate C_c and decreasing the rate C_i in certain amount does not change the cost too much. However, this observation is significant: customers are usually sensitive to the call interruption, and insensitive to the new call blocking, using discounts to the interrupted calls may help customers to reconciliate themselves; while this cost is compensated by increasing certain amount of the rate for complete calls, to which the customers are less sensitive. Thus, this scheme here is also put customer care into consideration. This planning is also easy to advertise and likely to be used in the future.

4.4.3 Nonlinear Charging Planning

In nonlinear charging planning, the charging rates for complete calls and incomplete calls are nonlinear functions of the effective holding times for complete calls and incomplete calls. The cost per call may be expressed as

$$C = \tilde{p}_i C_i(T_i) + \tilde{p}_c C_c(T_c), \tag{4.62}$$

where $C_i(\cdot)$ and $C_c(\cdot)$ are nonlinear functions. By choosing these two functions, we can have different rating planning. It is obvious, if we choose $C_i(x) = C_c(x) = C(x)$ and $C(x)$ is a convex function, then we have

$$C \geq C(\tilde{p}_i T_i + \tilde{p}_c T_c) = C(T),$$

which implies that the rating scheme using T_i and T_c separately can generate more revenue than the rating scheme using the single parameter T. The convexity of $C(\cdot)$ implies that we encourage medium length of calls, and discourage very short calls and very long calls by applying higher rate to those calls (observing the shape of convex functions). Another practical example is to choose the function $C_i(x)$ to be

$$C_i(x) = \begin{cases} \underline{c}_i & x \leq \underline{T}_i \\ c_i & \underline{T}_i < x \leq \overline{T}_i \\ \overline{c}_i & x > \overline{T}_i \end{cases}$$

while $C_c(\cdot)$ may choose linear function. In this example, we may choose the parameter $\underline{c}_i, c_i$ and $\overline{c}_i$ such that $c_i \leq \min\{\underline{c}_i, \overline{c}_i\}$. This reflects the fact that the setup procedure is more expensive while the longer calls are discouraged in order to accommodate more users. Customer care is an important issue in the future PCS networks, this scheme deserves much more attention. We will investigate this scheme in the future.

4.4.4 Adaptive Charging Planning

Customers may have different calling habits at different places at different times, service providers may try to determine best rating for their customers to show they care for their customers. According to the usage of a customer, the service providers may try to adaptively changing their charging plan to provide the customer an optimal rate (as a service provider it not only allures its customers to use its service, but more importantly keep them as its customer in order to stay competitive). By studying the schemes discussed above, the service providers can determine a rate to please its customers.

In summary, the effective call holding times for a complete call and an incomplete call are very important quantities in billing rate planning. More research in this area need to be done.

4.5 Illustrative Examples

In this section, we present a few illustrative examples to show how the results obtained in this chapter can be used to evaluate the performance of PCS networks. We will focus on the following scenarios:

(1). The cell residence times are iid according to the Gamma distribution with parameter $\gamma = 1.5$ and with different η values (change of mobility);

(2). The call holding times are iid according to the following distributions

- exponential distribution with parameter $\mu = 1/1.76$
- Gamma distribution with parameters $(2, \mu)$
- $r-$stage exponential distribution with parameters $r = 2$, $\mu_1 = 3\mu$ and $\mu_2 = 1.5\mu$
- hyper-exponential distribution with parameters $r = 2$, $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\mu_1 = 0.8\mu$ and $\mu_2 = 1.2\mu$
- hyper-Erlang distribution with parameters $r = 2$, $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $m_1 = 1$, $m_2 = 2$, $\mu_1 = 0.8\mu$ and $\mu_2 = 1.2\mu$

(3). The mobility η/μ is changing from 0 to 25.

In our examples, the considered distributions have the same expected values (for exponential and Gamma distributions, their expectations are $1/\mu$, for $r-$stage exponential distribution, its expectation is $1/\mu_1 + 1/\mu_2$, for hyper-exponential and hyper-Erlang distributions, their expectations are $\alpha_1/\mu_1 + \alpha_2/\mu_2$). The choice of μ is according to the wired telephony trials, where the expected call holding time is approximately 1.76 minutes ([34]).

Figures 4.4 and 4.5 show the call completion probability for different call holding time distributions in the above scenarios (the marking in Figure 4.5 is used in Figure 4.4 for curve reading). It can be observed the following:

- The call completion probability is insensitive to the call holding time distributions, especially for low mobility;
- The call completion probability is always decreasing as the mobility increases, which is consistent with our intuition that the higher the mobility, the more the handoffs, hence the higher the chance that the call will be incomplete, or the smaller the call completion probability.

Figures 4.6 and 4.7 show the expected effective call holding times for a complete call and an incomplete call, respectively. Although the type of the call holding time distributions does not significantly affect the call completion probability, it does greatly affect the effective call holding times for a complete call and an incomplete call. This is why we should consider the effective call holding times in evaluating the performance of PCS networks. From these two figures, we have the following observations:

- The expected effective call holding times (for either a complete call or an incomplete call) are always decreasing as mobility increases, as expected intuitively;
- For the expected effective call holding time of a complete call, the dependency on the type of call holding time distributions increases as the mobility increases. However, it is just the opposite for the expected effective call holding time of an incomplete call;
- The expected effective call holding time of a complete call is smaller than the absolute expected call holding time (the ideal case when there is no blocking and

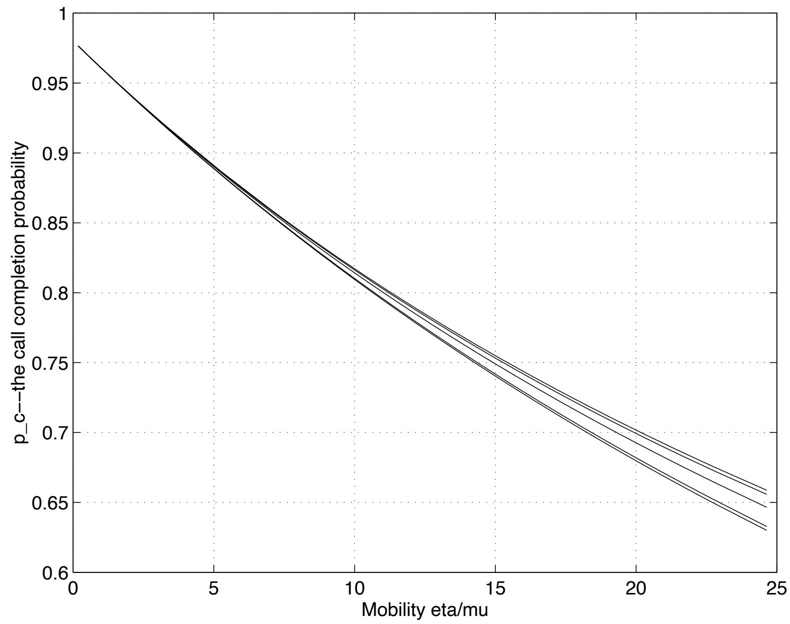


Figure 4.4: Call completion probability for different call holding time distributions

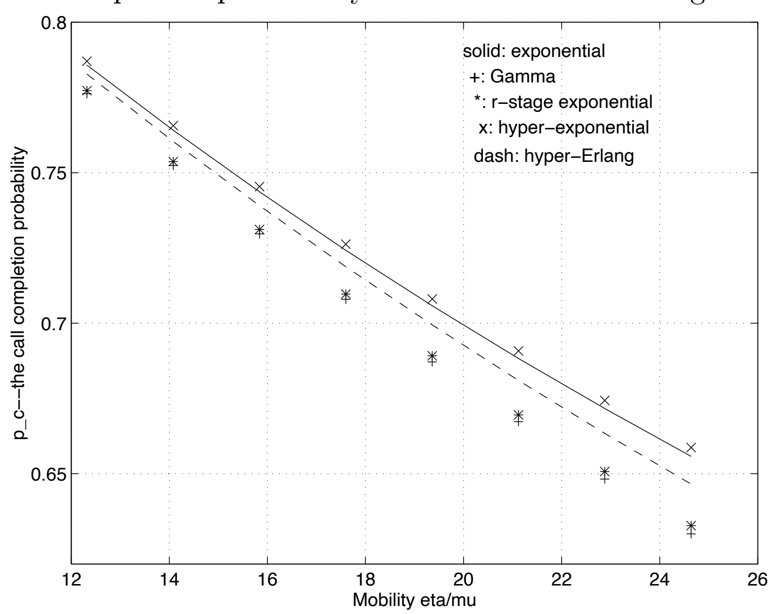


Figure 4.5: Call completion probability: localized piece of Figure 4.4

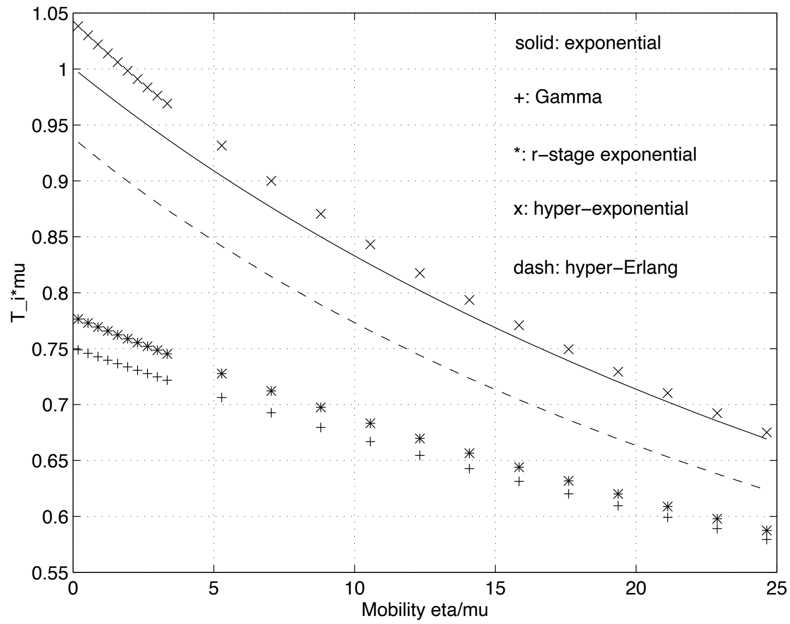


Figure 4.6: Expected effective call holding time for an incomplete call

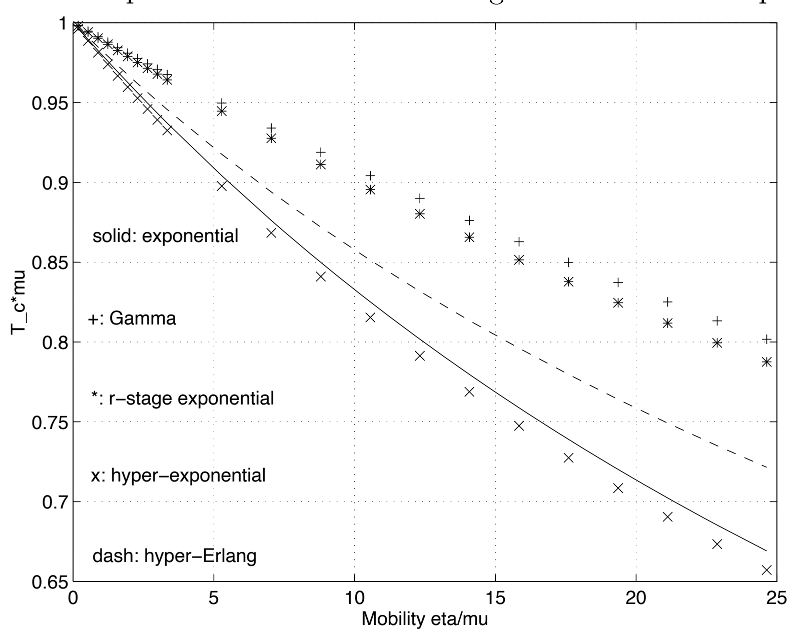


Figure 4.7: Expected effective call holding time for a complete call

no forced termination), i.e., $T_c \leq 1/\mu$, while the expected effective call holding time of an incomplete call can be larger than the absolute expected call holding time for ideal case, for example, $T_i > 1/\mu$ for hyper-exponentially distributed call holding times for low mobility.

4.6 Conclusions

In this chapter, we use a general distribution to model the call holding times and derive general formulae for the probability of a call completion and the expected effective call holding times of both complete and incomplete calls. By specifying the call holding time distribution to be Gamma, (staged) Erlang, hyperexponential and hyper-Erlang distributions, we obtain easy-to-compute formulae to compute these three quantities. These results can be expected to become significant in evaluating and tuning PCS network performance, and helpful in exercising the flat rate billing programs for these networks.

Chapter 5

BILLING STRATEGIES AND PERFORMANCE ANALYSIS

We now change to another topic: the billing strategies and performance analysis for the emerging PCS networks. Due to the fact that the roaming users may traverse a few registration areas in the process of a call, the classical billing systems for wired or cellular telephony may not be appropriate for the emerging PCS networks. Here we propose a few strategies specially for the roaming services.

5.1 Introduction

Billing World [49] predicts that by the year of 2000, US cellular carriers will invest US$1.6 billion for cellular billing and customer care. According to the managers from the top 20 cellular carriers, two of the most desirable attributes of the cellular billing systems are

- flexibility of upgrade and
- the capability to inform the in-house billing experts quickly about the status of

the system.

Telecommunication services are "culture sensitive", and the ways of service charging will significantly affect customer's behavior. For example, in many US cellular services, cellular subscribers are charged for the cellular usage, whether they are the calling or the called parties. Thus, cellular customers will tend to share their cellular phone numbers with only a small group of people to avoid "junk calls". On the other hand, the calling party is always charged for the cellular usage in some countries, for instance in Taiwan. Therefore, cellular subscribers (especially people using cellular phones for business) tend to distribute their cellular phone numbers as widely as possible to enhance their business opportunities. To maximize profits, a cellular carrier will therefore need to offer a variety of billing plans for same services, and may need change the plans from time to time to adapt to changing "customer culture". Thus, flexibility with upgrade is considered a highly important attribute of the billing system.

Another important attribute is the provision of quick billing status report, which is essential for the monitoring and diagnosing the billing system. One way to achieve this attribute is to report the customer billing records in real time. Unfortunately, this feature is not supported in most existing cellular billing systems. By comparison, in the *Public Switched Telephone Network* or *PSTN* (wireline network), real time billing information is possible. Typically, billing information will be delivered in the *Signaling System No. 7* (SS7) messages [41] during the call setup/release processes. The billing information is produced from the *automatic message accounting* (AMA) records. During the call setup/release processes, the monitor system tracks SS7 messages of the call and generates a *call detail record* (CDR) in AMA format when the call is completed. The CDR record will be stored in *Bellcore Accounting Format* (BAF), and the data can be transfered to the rating and billing systems.

The difficulty of providing real-time cellular customer billing records is due to the fact that the cellular users may roam from their "home systems"(HS) to the "visited systems"(VS). When a cellular user is in a visited system, the billing records for all call activities are kept in the visited system. In the existing cellular roaming management/call control protocols [37], there is no interaction between the visited system and the home system at the end of a call. Typically the billing information is kept in the visited system as a "roam" type *Cellular Intercarrier Billing Exchange Record* (CIBER). The roam CIBERs will be batched and periodically sent to a clearinghouse electronically or via mail in a tape format, and later forwarded by the clearinghouse to the customer's home system. The whole process may take from five days to more than two weeks. To speed up the billing information transmission, a cellular billing transmission standard called EIA/TIA IS-124 has been developed by working group four of TIA's TR 45.2 committee [2]. IS-124 will allow real-time billing information exchange, which will help control fraud by reducing the lag time created by the use of overnighted tape messages. Version A of IS-124 will also accommodate both US AMPS and international GSM carriers, which is desirable for heterogeneous PCS system integration [39].

An important performance issue of cellular billing information transmission is the frequency of the billing information exchange. In the ideal case, records would be transmitted for every phone call to achieve the real-time operation. However, real-time transmission would significantly increase the cellular signaling traffic, and seriously overload the signaling network of the PSTN. In order to achieve quick billing status report, a tradeoff is therefore needed between the frequency of the billing information transmission and the signaling traffic.

In this chapter, we propose three strategies for controlling the billing traffic for future PCS networks, which can be easily implemented in the billing systems of current cellular systems. The first strategy is to update the billing information when a fixed number of calls have been handled. The second strategy is to report the customer billing record in fixed time intervals. The third strategy is a combination of these two strategies, in which thresholds for both time and the number of calls are used.

We provide performance analysis of the various strategies, resulting in a set of guidelines, which can be used to select the appropriate frequency of transmissions for different network engineering requirements.

5.2 The Billing Checkpointing Models

The purpose of real-time billing information transmission is to ensure that when one (either a billing expert of the home system or a querying customer) checks the customer billing records, the information is up-to-date. The time of checking is referred to as the *billing information retrieval point* or *retrieval point.* Very often the billing record observer may tolerate certain degree of information obsoleteness (because the home system has not received the records of the outstanding calls) We define an observation at a retrieval point as *k-call outstanding* if the most recent k call records have not been received by the home system at the retrieval point. We can also define an observation at a retrieval point as $x-$*time outstanding* if the time elapsed at the retrieval point from the last record update is no more than x time units. It is very important to know the distributions of the number of the outstanding calls and the outstanding time. Since the billing system makes decisions based on the billing information received from the visited systems, the degree of "obsoleteness" of the billing information affects the accuracy of the billing decisions. Thus, it is desirable to know the average number of outstanding calls and the average total calling time of outstanding calls for the performance evaluation of billing systems.

We propose three strategies for the provision of quick billing status report while keeping the signaling traffic due to the on-line billing update in a certain tolerable level.

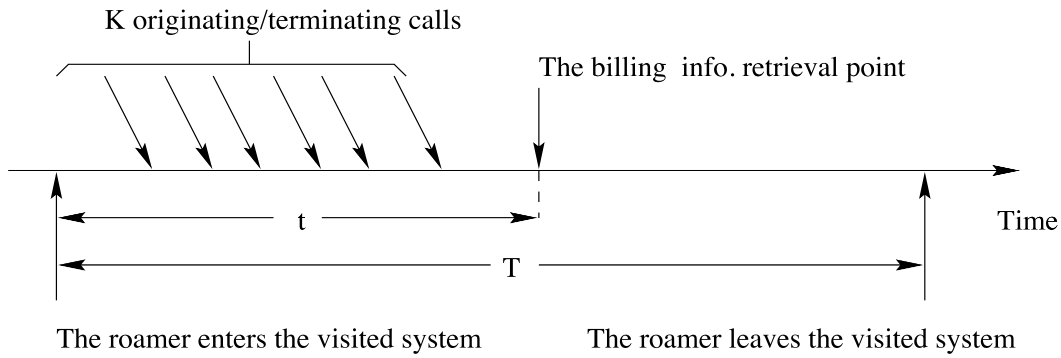


Figure 5.1: The timing diagram for the Checkpointing Model based on Number of Calls

5.2.1 Checkpointing Model Based on the Number of Calls

Assume that the billing information of a roamer is sent back to the home system for every n calls. We refer to n as the *checkpointing interval*, where $0 \leq k < n$. The real-time requirement in the cellular billing system specification may be "the probability of less than k-call outstanding observation should be larger than $\theta\%$". Thus, based on the mobility and call activities of a user, an appropriate n value can be selected to satisfy the above requirement.

Note that when the user leaves the visited system, the home system will send a *de-registration* or *cancellation* message to the visited system informing it that the user left the visited system, and the visited system will acknowledge the de-registration ([51]). We assume that the not-yet checkpointed billing records will be sent back to the home system through the acknowledgment, and no extra billing transmission message will be created.

Suppose that the roamer enters a visited system at time 0. The roamer resides at the visited system for a period T as shown in Figure 5.1. We call T as the *Visited System* (or VS) *residence time*. Suppose that the roamer's billing information is retrieved at the home system at time t, i.e., an expert or a customer of the home system queries the billing information at time t. If there are K phone calls to the roamer during $[0, t]$, then the billing retrieval is k-call outstanding if $K = in + k$ for some $i \geq 0$. We note that a billing record for a call is created when the call is completed. Thus, "call arrival time" in this context means that the time when the call record is created. Let $\Pr[k]$ be the probability that a billing retrieval is k-call outstanding.

Assume that T has a general density function $f(T)$ with the Laplace transform $f^*(s) = \int_{T=0}^{\infty} f(T)e^{-sT}dT$ and the mean $E[T] = 1/\eta$. Let the calls to the roamer be a Poisson process with arrival rate λ, and the billing retrieval point be a random observer. Let $r(t)$ and $r^*(s)$ be the density function and the Laplace Transform of the interval t in Figure 5.1. If $f(T)$ is non-lattice, then from the residual life theorem [47] we have

$$r(t) = \eta \int_{\tau=t}^{\infty} f(\tau)d\tau, \quad \text{and} \quad r^*(s) = \frac{\eta}{s}\Big[1 - f^*(s)\Big] \tag{5.1}$$

The billing retrieval is k-call outstanding if there are $K = in + k$ $(0 \leq k < n, i \geq 0)$ call arrivals during the period t. Since the call arrivals during the period t are a Poisson process, the k-call outstanding probability $\Pr[k]$ is expressed as

$$\begin{aligned}
\Pr[k] &= \int_{t=0}^{\infty} \sum_{i=0}^{\infty} \frac{(\lambda t)^{in+k}}{(in+k)!} e^{-\lambda t} r(t)dt \\
&= \sum_{i=0}^{\infty} \frac{\lambda^{in+k}}{(in+k)!} \int_{t=0}^{\infty} t^{in+k} r(t) e^{-\lambda t} dt \\
&= \sum_{i=0}^{\infty} \frac{(-\lambda)^{in+k}}{(in+k)!} \left[\frac{d^{in+k} r^*(s)}{ds^{in+k}}\right]\Bigg|_{s=\lambda}
\end{aligned} \tag{5.2}$$

From (5.1), (5.2) is re-written as

$$\begin{aligned}
\Pr[k] &= \sum_{i=0}^{\infty} \frac{\eta(-\lambda)^{in+k}}{(in+k)!} \left\{\frac{d^{in+k}}{ds^{in+k}}\left(\frac{1}{s}\right) - \frac{d^{in+k}}{ds^{in+k}}\left[\frac{f^*(s)}{s}\right]\right\}\Bigg|_{s=\lambda} \\
&= \sum_{i=0}^{\infty} \frac{\eta(-\lambda)^{in+k}}{(in+k)!} \Big[A - B\Big]\Bigg|_{s=\lambda}
\end{aligned} \tag{5.3}$$

where

$$A = \frac{d^{in+k}}{ds^{in+k}}\left(\frac{1}{s}\right) = -\frac{(in+k)!}{(-s)^{in+k+1}} \tag{5.4}$$

and

$$B = \frac{d^{in+k}}{ds^{in+k}} \left[\frac{f^*(s)}{s} \right] \tag{5.5}$$

For two functions $u(s)$ and $v(s)$ we have

$$\frac{d^n[u(s)v(s)]}{ds^n} = \sum_{j=0}^{n} \binom{n}{j} \left[\frac{d^j u(s)}{ds^j} \right] \left[\frac{d^{n-j} v(s)}{ds^{n-j}} \right]$$

let $u(s) = \frac{1}{s}$ and $v(s) = f^*(s)$, (5.5) is re-written as

$$B = -\sum_{j=0}^{in+k} \binom{in+k}{j} \left[\frac{j!}{(-s)^{j+1}} \right] \left[\frac{d^{in+k-j} f^*(s)}{ds^{in+k-j}} \right] \tag{5.6}$$

We use $u^{(i)}(s)$, in the subsequent development, to denote the $i-$th order derivative of function $u(s)$ at point s. From (5.3), (5.4) and (5.6), we have

$$\begin{aligned}
&\Pr[k] = \sum_{i=0}^{\infty} \frac{\eta(-1)^{in+k+1}\lambda^{in+k}}{(in+k)!} \\
\times\quad & \left\{ \frac{(in+k)!}{(-s)^{in+k+1}} - \sum_{j=0}^{in+k} \binom{in+k}{j} \left[\frac{j!}{(-s)^{j+1}} \right] \left[\frac{d^{in+k-j} f^*(s)}{ds^{in+k-j}} \right] \right\} \Bigg|_{s=\lambda} \\
=\quad & \sum_{i=0}^{\infty} \eta \left\{ \left(\frac{\lambda^{in+k}}{s^{in+k+1}} \right) - \sum_{j=0}^{in+k} (-1)^{in+k-j} \left[\frac{\lambda^{in+k}}{(in+k-j)!s^{j+1}} \right] \left[\frac{d^{in+k-j} f^*(s)}{ds^{in+k-j}} \right] \right\} \Bigg|_{s=\lambda} \\
=\quad & \sum_{i=0}^{\infty} \eta \left\{ \left(\frac{1}{\lambda} \right) - \frac{1}{\lambda} \sum_{j=0}^{in+k} (-1)^{in+k-j} \left[\frac{\lambda^{in+k-j}}{(in+k-j)!} \right] \left[\frac{d^{in+k-j} f^*(s)}{ds^{in+k-j}} \right] \Bigg|_{s=\lambda} \right\} \\
=\quad & \sum_{i=0}^{\infty} \eta \left\{ \left(\frac{1}{\lambda} \right) - \frac{1}{\lambda} \sum_{j=0}^{in+k} (-1)^{j} \left[\frac{\lambda^{j}}{j!} \right] f^{*(j)}(\lambda) \right\} \\
=\quad & \frac{\eta}{\lambda} \sum_{i=0}^{\infty} \left[1 - \sum_{j=0}^{in+k} \frac{(-\lambda)^j}{j!} f^{*(j)}(\lambda) \right].
\end{aligned}$$

Thus, we have

Theorem 5.1.

$$\Pr[k] = \frac{\eta}{\lambda} \sum_{i=0}^{\infty} \left[1 - \sum_{j=0}^{in+k} \frac{(-\lambda)^j}{j!} f^{*(j)}(\lambda) \right].$$

From this theorem, we can develop an algorithm to compute the probability $\Pr[k]$. Since $f^*(s)$ is an analytic function at $s = \lambda$, hence

$$f^*(s) = \sum_{j=0}^{\infty} \frac{f^{*(j)}(\lambda)}{j!}(s-\lambda)^j, \quad 1 = f^*(0) = \sum_{j=0}^{\infty} \frac{f^{*(j)}(\lambda)}{j!}(-\lambda)^j,$$

hence from (5.7), we obtain

$$\Pr[k] = \frac{\eta}{\lambda}\sum_{i=0}^{\infty}\sum_{j=in+k+1}^{\infty}\frac{(-\lambda)^j}{j!}f^{*(j)}(\lambda) = \frac{\eta}{\lambda}\sum_{i=0}^{\infty}(i+1)\left(\sum_{j=in+k+1}^{(i+1)n+k}\frac{(-\lambda)^j}{j!}f^{*(j)}(\lambda)\right). \quad (5.7)$$

Thus, we first expand the $f^*(s)$ at $s = \lambda$ as Taylor series, using this to express $f^*(0)$ as $f^*(0) = \sum_{i=0}^{\infty}\alpha_j$, then using the sequence $\{\alpha_j\}$ and (5.7) to compute $w_i = \sum_{j=in+k+1}^{(i+1)n+k}\alpha_j$, hence $\Pr[k]$ can be computed by $\Pr[k] = \sum_{i=0}^{\infty}(i+1)w_i$. This algorithm is numerically easy to implement.

Let N_o denote the number of outstanding calls, i.e., the number of arriving calls during the interval between the last billing update and the retrieval point, the records of these calls are not available at the retrieval point at home system. Let t_{tot} be the total calling time of these outstanding calls, and let the arriving calls have the expected call holding time $1/\mu$. Then the expected number of outstanding calls is given by

$$E[N_o] = \sum_{k=0}^{n-1} k\Pr[k].$$

Let t_i denote the call holding time of an arriving call, then the total calling time of outstanding calls is given by

$$t_{tot} = \sum_{i=1}^{N_o} t_i.$$

From Wald's equation ([63]), we obtain

$$E[t_{tot}] = E[t_i]E[N_o] = E[N_o]/\mu.$$

Remark: If $1/\mu$ is interpreted as the average call holding times for possible fraudulent calls, then the maximum fraudulent calling times is $E[t_{tot}]$.

For special VS residence time distributions, analytical expression may be obtained. For example, if the VS residence times have an Erlang distribution with the shape parameter m and the scale parameter $\alpha = m\eta$ (and thus the variance of the distribution is $V = 1/(m\mu)^2$), then

$$f(t) = \frac{(\alpha t)^{m-1}}{(m-1)!}\alpha e^{-\alpha t}, \quad \text{and} \quad f^*(s) = \left(\frac{\alpha}{s+\alpha}\right)^m \quad (5.8)$$

Denote $\Pr_{(m,\alpha)}[k]$ as the probability $\Pr[k]$ when the Erlang VS residence distribution have the shape parameter m and the scale parameter α. Substitute (5.8) in (5.7) and after careful rearrangement using the combinatorial identity techniques [62], we have

$$\Pr_{(m,\alpha)}[k] = \sum_{i=0}^{\infty}\sum_{l=0}^{m-1} \eta \binom{in+k+l}{l} \frac{\lambda^{in+k}\alpha^{l}}{(\lambda+\alpha)^{in+k+l+1}} \tag{5.9}$$

(5.9) can be presented in a recursive format:

$$\Pr_{(m,\alpha)}[k] = \begin{cases} \sum_{i=0}^{\infty} \eta \binom{in+k+m-1}{m-1} \frac{\lambda^{in+k}\alpha^{m-1}}{(\lambda+\alpha)^{in+k+m}} + \Pr_{(m-1,\alpha)}[k], & \text{for } m \geq 1 \\ 0, & \text{for } m = 0 \end{cases} \tag{5.10}$$

Thus, we have

Corollary 5.1. The probability, $\Pr_{(m,\alpha)}[k]$, of k-call outstanding billing retrieval for Erlang (with parameters (m,α)) VS residence time is given by

$$\Pr_{(m,\alpha)}[k] = \begin{cases} \sum_{i=0}^{\infty} \eta \binom{in+k+m-1}{m-1} \frac{\lambda^{in+k}\alpha^{m-1}}{(\lambda+\alpha)^{in+k+m}} + \Pr_{(m-1,\alpha)}[k], & \text{for } m \geq 1 \\ 0, & \text{for } m = 0 \end{cases}$$

For $m = 1$ we have

$$\begin{aligned} \Pr_{(1,\alpha)}[k] &= \sum_{i=0}^{\infty} \frac{\eta\lambda^{in+k}}{(\lambda+\alpha)^{in+k+1}} \\ &= \left(\frac{\eta}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^{k} \left\{\sum_{i=0}^{\infty}\left[\left(\frac{\lambda}{\lambda+\alpha}\right)^{n}\right]^{i}\right\} \\ &= \left(\frac{\eta}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^{k}\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^{n}\right]^{-1}, \quad (\alpha=\eta). \end{aligned} \tag{5.11}$$

For $m = 2$ we have

$$\begin{aligned} \Pr_{(2,\alpha)}[k] &= \sum_{i=0}^{\infty} \frac{\eta\alpha(in+k+1)\lambda^{in+k}}{(\lambda+\alpha)^{in+k+2}} + \Pr_{(1,\alpha)}[k] \\ &= \sum_{i=0}^{\infty} \frac{in\eta\alpha\lambda^{in+k}}{(\lambda+\alpha)^{in+k+2}} + \sum_{i=0}^{\infty} \frac{(k+1)\eta\alpha\lambda^{in+k}}{(\lambda+\alpha)^{in+k+2}} + \Pr_{(1,\alpha)}[k] \end{aligned} \tag{5.12}$$

Since

$$\begin{aligned}\sum_{i=0}^{\infty} \frac{in\eta\alpha\lambda^{in+k}}{(\lambda+\alpha)^{in+k+2}} &= \frac{n\eta\alpha\lambda^k}{(\lambda+\alpha)^{k+2}} \left\{\sum_{i=0}^{\infty} i \left[\left(\frac{\lambda}{\lambda+\alpha}\right)\right]^i\right\} \\ &= \left[\frac{n\eta\alpha\lambda^k}{(\lambda+\alpha)^{k+2}}\right]\left(\frac{\lambda}{\lambda+\alpha}\right)^n \left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-2} \\ &= \left(\frac{n\alpha}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^n \left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1} \Pr_{(1,\alpha)}[k]\end{aligned}$$

and

$$\begin{aligned}\sum_{i=0}^{\infty} \frac{(k+1)\eta\alpha\lambda^{in+k}}{(\lambda+\alpha)^{in+k+2}} &= \frac{(k+1)\eta\alpha\lambda^k}{(\lambda+\alpha)^{k+2}} \left\{\sum_{i=0}^{\infty} \left[\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^i\right\} \\ &= \left[\frac{(k+1)\eta\alpha\lambda^k}{(\lambda+\alpha)^{k+2}}\right]\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1} \\ &= \left[\frac{(k+1)\alpha}{\lambda+\alpha}\right]\Pr_{(1,\alpha)}[k]\end{aligned}$$

(5.12) is rewritten as

$$\begin{aligned}\Pr_{(2,\alpha)}[k] &= \left\{\left(\frac{n\alpha}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^n\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1} + \left[\frac{(k+1)\alpha}{\lambda+\alpha}\right]+1\right\}\Pr_{(1,\alpha)}[k], \\ &\quad (\alpha = 2\eta). \qquad (5.13)\end{aligned}$$

For $m = 3$,

$$\begin{aligned}\Pr_{(3,\alpha)}[k] &= \sum_{i=0}^{\infty} \eta(in+k+2)(in+k+1)\left[\frac{\lambda^{in+k}\alpha^2}{(\lambda+\alpha)^{in+k+3}}\right] + \Pr_{(2,\alpha)}[k] \\ &= \left[\frac{\eta\lambda^k\alpha^2}{(\lambda+\alpha)^{k+3}}\right]\sum_{i=0}^{\infty}[i^2n^2+(2k+3)in+(k+2)(k+1)]\left[\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^i \\ &+ \Pr_{(2,\alpha)}[k] \qquad (5.14)\end{aligned}$$

Since

$$\sum_{i=0}^{\infty} i^2x^i = \frac{x(1+x)}{(1-x)^3}$$

(5.14) is re-written as

$$Pr_{(3,\alpha)}[k] = \frac{\eta\lambda^k\alpha^2}{(\lambda+\alpha)^{k+3}}$$

$$
\begin{aligned}
&\times \quad \left\{ n^2 \left(\frac{\lambda}{\lambda+\alpha}\right)^n \left[1+\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-3} \right.\\
&+ \quad n(2k+3)\left(\frac{\lambda}{\lambda+\alpha}\right)^n\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-2}\\
&+ \quad \left.(k+2)(k+1)\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1}\right\} + \Pr_{(2,\alpha)}[k], \quad (\alpha = 3\alpha). \quad (5.15)
\end{aligned}
$$

Summarizing above discussion, we obtain

Corollary 5.2.

$$
\begin{aligned}
\Pr_{(1,\alpha)}[k] &= \left(\frac{\eta}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^k\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1}, \quad (\alpha=\eta)\\
\Pr_{(2,\alpha)}[k] &= \left\{\left(\frac{n\alpha}{\lambda+\alpha}\right)\left(\frac{\lambda}{\lambda+\alpha}\right)^n\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1} + \left[\frac{(k+1)\alpha}{\lambda+\alpha}\right]+1\right\}\Pr_{(1,\alpha)}[k],\\
&\quad (\alpha = 2\eta)\\
Pr_{(3,\alpha)}[k] &= \frac{\eta\lambda^k\alpha^2}{(\lambda+\alpha)^{k+3}}\\
&\times \quad \left\{ n^2 \left(\frac{\lambda}{\lambda+\alpha}\right)^n \left[1+\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-3} \right.\\
&+ \quad n(2k+3)\left(\frac{\lambda}{\lambda+\alpha}\right)^n\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-2}\\
&+ \quad \left.(k+2)(k+1)\left[1-\left(\frac{\lambda}{\lambda+\alpha}\right)^n\right]^{-1}\right\} + \Pr_{(2,\alpha)}[k], \quad (\alpha = 3\eta).
\end{aligned}
$$

It is observed that when η is sufficiently small, i.e., the roamer stays in that visited area for sufficiently long time, the number of outstanding calls are equally probable, i.e., uniformly distributed. Indeed, we can easily show that

$$
\lim_{\eta\to 0}\Pr_{(1,\alpha)}[k] = \frac{1}{n}, \quad \lim_{\eta\to 0}\Pr_{(2,\alpha)}[k] = \frac{1}{n}.
$$

We conjecture that this observation is valid for any nonlattice distribution $f(T)$.

When the VS residence time is exponentially distributed, i.e., $m = 1$, we have

$$
E[N_o] \;=\; \sum_{k=0}^{n-1} k\Pr[k] = \sum_{k=0}^{n-1}\left(\frac{\eta}{\lambda+\eta}\right)\left(\frac{\lambda}{\lambda+\eta}\right)^k\left[1-\left(\frac{\lambda}{\lambda+\eta}\right)^n\right]^{-1}
$$

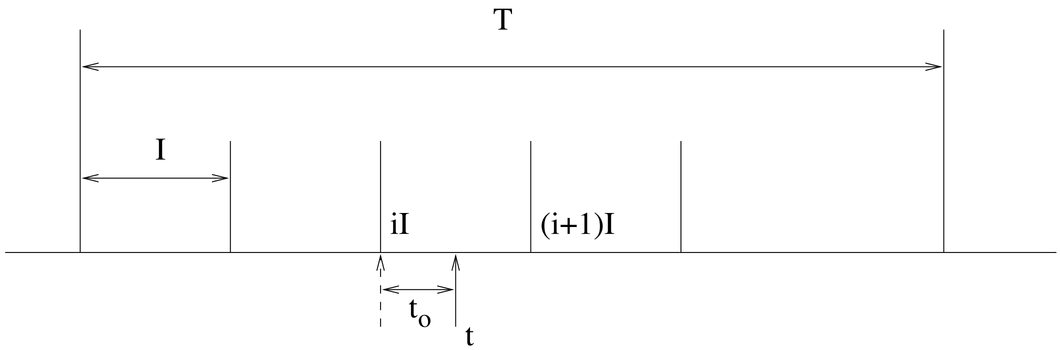


Figure 5.2: Time Diagram for Checkpointing Model based on Time Interval

$$= \left(\frac{\eta}{\lambda+\eta}\right)\left[1-\left(\frac{\lambda}{\lambda+\eta}\right)^n\right]^{-1}\left(\frac{\lambda}{\lambda+\eta}\right)\sum_{k=1}^{k-1} k\left(\frac{\lambda}{\lambda+\eta}\right)^{k-1}$$
$$= \frac{\lambda}{\eta}\left[1-\left(\frac{\lambda}{\lambda+\eta}\right)^n\right]^{-1}\left[1-n\left(\frac{\lambda}{\lambda+\eta}\right)^{n-1}+(n-1)\left(\frac{\lambda}{\lambda+\eta}\right)^n\right],$$

and

$$E[t_{tot}] = \frac{\lambda}{\eta\mu}\left[1-\left(\frac{\lambda}{\lambda+\eta}\right)^n\right]^{-1}\left[1-n\left(\frac{\lambda}{\lambda+\eta}\right)^{n-1}+(n-1)\left(\frac{\lambda}{\lambda+\eta}\right)^n\right].$$

5.2.2 Checkpointing Model Based on Time Interval

In the previous subsection, we applied a threshold to the number of calls in the update of billing status. This approach may have a disadvantage when the call holding times are long. For example, when a fraudulent user, using another customer's identification, makes a call to a 900 number in the US, the call holding time tends to be quite long. During a typical billing period, the number of such calls may not be high, therefore the previous strategy may not be appropriate. One solution to overcome this is to use time interval (threshold in time) for updating of billing record. In this subsection, we analyze this strategy.

Assume now that the billing information is sent back to the home system (HS) from the visited system (VS) for every I time units. As in the previous section, we assume that the VS residence time T has a nonlattice density function $f(T)$ with

Laplace transform $f^*(s)$ and with expected residence time $1/\eta$. Let $F(T)$ be its distribution function. Let t be the retrieval point. Let t_o denote the outstanding time, i.e., the time between the retrieval point and the last update instant. Since the customer billing record is forwarded to the HS every I time units, t_o is distributed in the interval $[0, I]$. Let $r(t)$ and $r^*(s)$ be the density function and Laplace transform of the interval t. As above, I is the updating interval, i.e., the billing information will be forwarded from the VS to HS every I time units whenever the roamer is still in that visited area. Using the Residual Life Theorem ([47]) we have

$$r(t) = \eta(1 - F(t)), \quad \text{and} \quad r^*(s) = \frac{\eta}{s}(1 - f^*(s)). \tag{5.16}$$

As shown in Figure 5.2, t_o is the time between the last updating instant, say, iI and the checking point t, so

$$t_o = t - iI = t(mod\ I),$$

where $(mod\ I)$ here means the residue modulo I. Let $f_o(t)$ and $F_o(t)$ denote the density function and distribution function of t_o, respectively. Then for any $0 \leq x \leq I$, we have

$$\begin{aligned}
F_o(t) &= \Pr(t_o \leq x) = \Pr(t(mod\ I) \leq x) \\
&= \sum_{i=0}^{\infty} \Pr(t(mod\ I) \leq x,\ iI \leq t \leq (i+1)I) \\
&= \sum_{i=0}^{\infty} \Pr(t - iI \leq x, iI \leq t \leq (i+1)I) \\
&= \sum_{i=0}^{\infty} \Pr(iI \leq t \leq iI + x) \\
&= \sum_{i=0}^{\infty} \int_{iI}^{iI+x} r(t)dt = \sum_{i=0}^{\infty} \eta \int_{iI}^{iI+x} (1 - F(t))dt
\end{aligned}$$

so the density function of t_o is given by

$$f_o(x) = \sum_{i=0}^{\infty} r(iI + x) = \eta \sum_{i=0}^{\infty} [1 - F(iI + x)]. \tag{5.17}$$

Suppose that the Laplace transform of the visited system residence time $f^*(s)$ has only isolated singular points in the open left half complex plane, let σ_p be the set of poles of $f^*(s)$. Let $\sigma < 0$ be such real number such that all poles have real parts smaller than σ. It is obvious that $r^*(s)$ and $f^*(s)$ have the same poles. From the inverse Laplace transform formula and the Residue Theorem ([35]), we have (the contour used in the Residue Theorem is chosen to be the domain in the left half plane to the left of vertical line $\mathrm{Re}(s) = \sigma$)

$$\begin{aligned} f_o(x) &= \sum_{i=0}^{\infty} \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} r^*(s) e^{s(iI+x)} ds \\ &= \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} r^*(s) \left(\sum_{0}^{\infty} e^{iIs} \right) e^{sx} ds \\ &= \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{r^*(s) e^{sx}}{1 - e^{sI}} ds = \frac{1}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} \frac{\eta(1 - f^*(s)) e^{sx}}{s(1 - e^{sI})} ds \\ &= \sum_{p \in \sigma_p} \operatorname*{Res}_{s=p} \frac{\eta(1 - f^*(s)) e^{sx}}{s(1 - e^{sI})} = -\eta \sum_{p \in \sigma_p} \operatorname*{Res}_{s=p} \frac{f^*(s) e^{sx}}{s(1 - e^{sI})}, \end{aligned} \tag{5.18}$$

where $\mathrm{Res}_{s=p}$ denotes the residue at the pole $s = p$.

Assume that the call arrivals for the roamer is Poisson with arrival rate λ. Let N_o denote the number of outstanding calls (the number of calls arriving after the last billing update instant and before the retrieval point, the billing records of these outstanding calls are not available at the checking point), then the probability that there are k outstanding calls is given by

$$\Pr(N_o = k) = \int_0^I \Pr(N_o = k | t_o = x) f_o(x) dx = \int_0^I \frac{(\lambda x)^k}{k!} e^{-\lambda x} f_o(x) dx. \tag{5.19}$$

Following a similar procedure as in the derivation of (5.18), the expected number of outstanding calls is given by

$$\begin{aligned} & E(N_o) = \sum_{k=0}^{\infty} k \Pr(N_o = k) = \int_0^I \left(k \frac{(\lambda x)^k}{k!} \right) e^{-\lambda x} f_o(x) ds \\ = & \int_0^I \lambda x f_o(x) dx = \lambda E(t_o) = \lambda \sum_{i=0}^{\infty} \int_0^I x r(iI + x) dx \\ = & \frac{\lambda}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} r^*(s) \int_0^I x e^{sx} \left(\sum_{i=0}^{\infty} e^{isI} \right) dx ds \\ = & \frac{\lambda}{2\pi j} \int_{\sigma - j\infty}^{\sigma + j\infty} r^*(s) \int_0^I \frac{x e^{sx}}{1 - e^{sI}} dx ds \end{aligned}$$

$$
\begin{aligned}
&= \frac{\lambda}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{r^*(s)[(-1+sI)e^{sI}+1]}{s^2(1-e^{sI})}ds \\
&= \frac{\lambda}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty}\frac{\eta(1-f^*(s))[(-1+sI)e^{sI}+1]}{s^3(1-e^{sI})}ds \\
&= \lambda\eta\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{(1-f^*(s))[(-1+sI)e^{sI}+1]}{s^3(1-e^{sI})} \\
&= -\lambda\eta\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{[1+(sI-1)e^{sI}]f^*(s)}{s^3(1-e^{sI})}. \qquad (5.20)
\end{aligned}
$$

Let $t_1, t_2, \ldots, t_k, \cdots$ be the call holding times of the roamer, assume that this random sequence is independently identically distributed with expected value $1/\mu$, then the total outstanding calling time (the summation of total calling times for all outstanding calls) is given by

$$
t_{tot} = \sum_{i=1}^{N_0} t_i. \qquad (5.21)
$$

From Wald's equation ([63]) we obtain the expected total outstanding time is given by

$$
E(t_{tot}) = E(N_o)E(t_i) = -\frac{\lambda\eta}{\mu}\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{[1+(sI-1)e^{sI}]f^*(s)}{s^3(1-e^{sI})}. \qquad (5.22)
$$

Summarizing above, we obtain

Theorem 5.2. Suppose that $f^*(s)$ has only isolated singular points, let σ_p denote the set of poles of $f^*(s)$. Let $\mathrm{Res}_{s=p}$ denote the residue operator at the pole $s = p$ ([35]). Then, $f_o(x)$ can be expressed as follows:

$$
\begin{aligned}
f_o(x) &= -\eta\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{f^*(s)e^{sx}}{s(1-e^{sI})}; \\
E[N_o] &= -\lambda\eta\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{[1+(sI-1)e^{sI}]f^*(s)}{s^3(1-e^{sI})}; \\
E[t_{tot}] &= -\frac{\lambda\eta}{\mu}\sum_{p\in\sigma_p}\operatorname*{Res}_{s=p}\frac{[1+(sI-1)e^{sI}]f^*(s)}{s^3(1-e^{sI})}.
\end{aligned}
$$

For simplicity, let

$$
g(s;x) = \frac{e^{sx}}{s(1-e^{sI})}, \quad h(s) = \frac{1+(sI-1)e^{sI}}{s^3(1-e^{sI})}. \qquad (5.23)
$$

If the VS residence times have an Erlang distribution with the shape parameter m and the scale parameter $\alpha = m\eta$, then

$$f^*(s) = \left(\frac{\alpha}{s+\alpha}\right)^m.$$

This function has only one pole $s = -\alpha$, i.e., $\sigma_p = \{-\alpha\}$. From Theorem 5.2, we obtain

Corollary 5.3. If the VS residence times are Erlang distributed with parameter (m, α), then

$$f_o(x) = -\frac{\eta\alpha^m}{(m-1)!}\frac{\partial^{m-1}}{\partial s^{m-1}}g(s;x)\bigg|_{s=-\alpha}, \tag{5.24}$$

$$E(N_o) = -\frac{\lambda\eta\alpha^m}{(m-1)!}h^{(m-1)}(-\alpha), \tag{5.25}$$

$$E(t_{tot}) = -\frac{\lambda\eta\alpha^m}{(m-1)!\mu}h^{(m-1)}(-\alpha). \tag{5.26}$$

In particular, when $m = 1$, i.e., the VS residence times are exponentially distributed, we have

$$\begin{aligned} f_o(x) &= \frac{\eta e^{-\eta x}}{1-e^{-\eta I}}, \\ E(N_o) &= \frac{\lambda}{\eta}\cdot\frac{1-(1+\eta I)e^{-\eta I}}{1-e^{-\eta I}}, \\ E(t_{tot}) &= \frac{\lambda}{\eta\mu}\cdot\frac{1-(1+\eta I)e^{-\eta I}}{1-e^{-\eta I}}. \end{aligned}$$

When $m = 2$, the outstanding time distribution is given by

$$\begin{aligned} f_o(x) &= -\eta(2\eta)^2\frac{\partial}{\partial s}g(s;x)\bigg|_{s=-2\eta} \\ &= \frac{\eta\left[(1+2\eta x)(1-e^{-2\eta I})+2\eta I e^{-2\eta I}\right]}{(1-e^{-2\eta I})^2}e^{-2\eta x}. \end{aligned}$$

The expected number of outstanding calls is given by

$$\begin{aligned} & E(N_o) = -\lambda\eta(2\eta)^2 h^{(1)}(-2\eta) \\ = \;& \frac{\lambda\left\{\left[1-(1+2\eta I)e^{-2\eta I}\right]\left[3(1-e^{-2\eta I})+2\eta I e^{-2\eta I}\right]-(2\eta I)^2(1-e^{-2\eta I})e^{-2\eta I}\right\}}{4\eta(1-e^{-2\eta I})^2}. \end{aligned}$$

And the total outstanding calling time is given by

$$E(t_{tot}) = \frac{\lambda\left\{\left[1-(1+2\eta I)e^{-2\eta I}\right]\left[3(1-e^{-2\eta I})+2\eta I e^{-2\eta I}\right]-(2\eta I)^2(1-e^{-2\eta I})e^{-2\eta I}\right\}}{4\eta\mu(1-e^{-2\eta I})^2}.$$

It is not difficult to show that when η is sufficiently small, the density function $f_o(x)$ for either $m=1$ or $m=2$ approaches to $1/I$, which implies that the outstanding time is uniformly distributed when the roamer stays in the visited area for sufficiently long. We conjecture that this is true for any nonlattice distributed VS residence time.

It is obvious that we can apply different distributions such as Gamma, hyper-exponential and hyper-Erlang distribution to model the VS residence times and similar analytic results can be obtained. This is left to the readers as exercises.

5.2.3 Checkpointing Model based on Number of Calls and Time Interval

The last strategy uses the time threshold for updating billing record. This strategy uses both the number of calls and time interval for updating billing record. The thresholds used in this case (n and I) may be less restricted (larger) than those in previous strategies. In this strategy, the billing record update works as follows: Choose threshold n for the number of calls and threshold I for the time interval. The billing record at the VS keeps the records of both time, say, t_3 and the number of calls, say, n_3. When a customer enters a visited area, the billing record for it starts, and set $t_3 = 0$ and $n_3 = 0$, the clock for t_3 starts. If there are calls from or to the customer, n_3 is incremented by one. Whenever either t_3 or n_3 reaches the threshold, the billing record is forwarded to the HS, at the same time t_3 and n_3 are reset to zeros, and starts over again if the customer is still in the VS area. The methods used in previous two sections can be modified to analyze the performance of this model.

5.2.4 Choosing the Threshold

In the above strategies, we need to determine how often the customer's billing record should be updated, i.e., how to choose the thresholds, since if the billing record is forwarded from VS to the HS one by one in real time, the signaling channel may be congested.

To this end we assume that the call arrivals of a customer have Poisson distribution with parameter λ, let $\tau_1, \tau_2, \ldots$ denote the interarrival times whose expected value is $1/\lambda$.

We consider the first strategy. In this strategy, the number of outstanding calls is related to the call arrival process. Suppose that the billing system may tolerate the outstandingness of time for a period of $T_h > 0$, which is determined by the PCS network design objective and the availability of the bandwidth for signaling channel. Then the total time that n call arrives is given by

$$T_{tot} = \tau_1 + \tau_2 + \cdots + \tau_n,$$

we need $E[T_{tot}] \leq T_h$, which is equivalent to

$$n \leq \lambda T_h.$$

So the threshold can be chosen as the integral part of the positive number λT_h.

In the second strategy, the billing system will be tolerant of a fixed number of outstanding calls. This is interpreted as the *probability of more than, say, K calls in the interval of length I is less than a preassigned number, say, θ ($0 < \theta < 1$).* Suppose that there are N calls in the updating interval with length I, then the above criterion can be translated into the following

$$Pr(N > K) = \sum_{i=K+1}^{\infty} Pr(N = i) = \sum_{i=K+1}^{\infty} \frac{(\lambda I)^i}{i!} e^{-\lambda I} < \theta,$$

i.e.,

$$p(I) \stackrel{\text{def}}{=} \sum_{i=0}^{K} \frac{(\lambda I)^i}{i!} e^{-\lambda I} \geq 1 - \theta. \tag{5.27}$$

Notice that since $\frac{dp(I)}{dI} = -\frac{(\lambda I)^K}{K!}\lambda e^{-\lambda I} < 0$ for any $I > 0$, so $p(I)$ is a strictly decreasing function of I. Also, $p(0) = 1$ and $p(+\infty) = 0$, then for any $0 \leq \theta < 1$, we can always find solution for (5.27). For very small T, (5.27) is always true, but the small value for I may not be desirable due to the channel traffic intensity requirement. The best value would be the largest I that satisfies (5.27). It is obvious that there is a unique solution for this value, which is denoted by $T_{\max}$. We can choose $I = T_{\max}$ in the second strategy. The thresholds in the third strategy can use either (T_h, K), a conservative choice, or (I, n) from the previous two strategies based on the value (T_h, K).

5.2.5 On-line Detection of Fraudulent Activities

The previous billing strategies can be used for on-line detection of fraudulent activities. In PCS networks, the calling record for a user is usually stored in the HS. When the user roams to another area and uses his portable through the PCS network, its calling record will temporarily be stored in the VS, then after some time (possibly hours, days or weeks) the collected calling record will be forwarded to the HS for accounting and billing purposes. A fraudulent activity may happen during this transaction time, when an invader uses another customer's identification to make connections to the PCS network. This is particularly true for analog cellular systems where fraud is possible through cloning of the portable. The user, or the system operator, can find out about the fraud only after reviewing their calling record when it becomes available, i.e. after the damage has been already done.

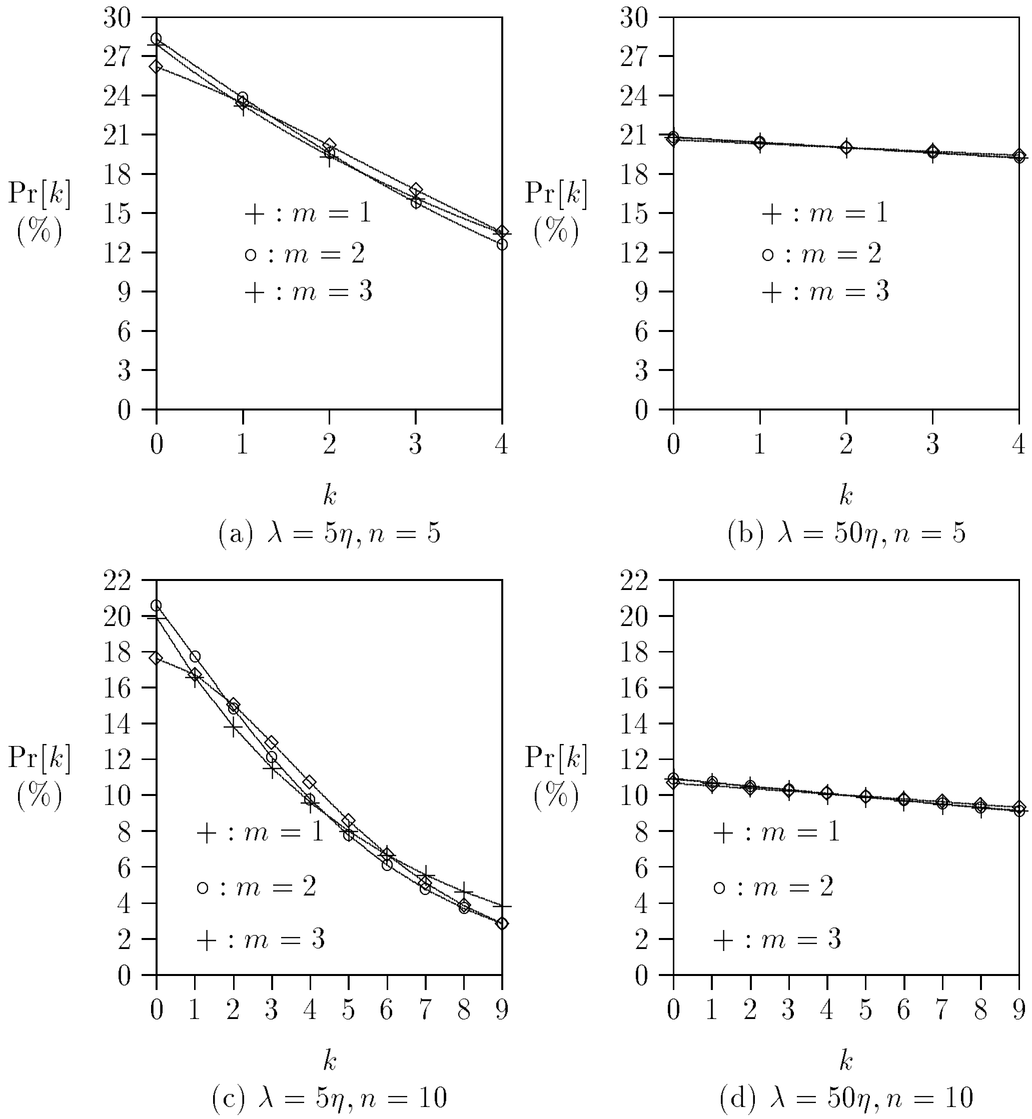
With the our proposed strategies, we can implement a model to detect some of the fraudulent activities more effectively than the classical billing systems where the fraudulent activity is detected only after the hard copy of the billing record is delivered to the main office after possible long time. In our model, when the billing record status report is transferred to the HS through signal channels, at each update or a few updates, a statistical analysis is applied to the quantities which characterize the user's calling usage. If these quantities do not conform to the user's registered usage, an alarm signal will be sent out, and the operator may take actions. For instance, he can initiate a call to the user to verify whether the user has made those calls. In this way, the possible fraud can be detected in one updating period, thus significantly reducing the cost of fraudulent activity.

The quantities used for this evaluation may be mean and variance of calling time or the number of calls in a certain period of time, etc., in conjunction with certain thresholds (for alarm signals) of the statistical quantities, for the statistical decision making. This statistical analyzer can be easily implemented at the HS.

One way to further reduce the signaling traffic for the purpose of fraud detection is to reduce the data contained in the billing record. A simplified billing record can be used. Such simplified billing record may contain just the phone number and the call holding time for each call without other details. In this way, the thresholds n and I can be smaller than those used in the actual billing updates. The smaller thresholds can be useful for the approximation of real-time billing.

5.3 Illustrative Examples

We first consider the model based on the number of outstanding calls. Based on Corollary 5.2, Figure 5.3 illustrates the effect of the Erlang shape parameter m on the k-call outstanding probability $\Pr_{(m,\alpha)}[k]$. In the figures, the mean VS residence time is $E[T] = \frac{1}{\eta}$. $E[T]$ can be several hours, days, or months. For simplicity, the call arrival rate λ is normalized by η in our numerical examples. Each figure plots the $\Pr[k]$ curves for $m = 1, 2$, and 3 (i.e., $\alpha = \eta, 2\eta$, and 3η), $n = 5, 10$, and $\lambda = 5\eta, 50\eta$. That is, we consider the Erlang VS residence times T with the same mean $1/\eta$, but different variances $1/\eta^2$, $1/(2\eta^2)$, and $1/(3\eta^2)$, respectively. The figures indicate that the shape parameter (or the variance) of the Erlang VS residence times do not have



(a) $\lambda = 5\eta, n = 5$

(b) $\lambda = 50\eta, n = 5$

(c) $\lambda = 5\eta, n = 10$

(d) $\lambda = 50\eta, n = 10$

Figure 5.3: Effects of m on $\Pr[k]$

significant effect on the k-call outstanding probability. This result is true for different n values and for large λ values. Figures 5.3 and 5.5 plots $\Pr[k]$ for different λ values, where $m = 1$, and $n = 5$ and 10, respectively. These figures indicate that

$$\Pr[k] > \Pr[k+1], \quad \text{for} \quad 0 \le k < n-1 \tag{5.28}$$

Let $\rho = \lambda/\eta$ be the call to mobility ratio. When $\rho < 3n$, $\Pr[k]$ may significantly larger than $\Pr[k+1]$. On the other hand, if $\rho >> n$, then $\Pr[k] \simeq \Pr[k+1]$. This phenomenon is explained as follows. In the period T, the last checkpointing is performed when the roamer moves out of the visited system, and no more than n call records are sent back to the home system. Thus, if a billing retrieval falls in this last checkpointing period, it is likely to have a small k-call outstanding observation. For the other "normal" n-checkpointing intervals, $\Pr[k]$ tends to be uniformly distributed (because the billing retrieval is a random observer). When $\rho >> n$, the end effect of the last checkpointing interval becomes insignificant and $\Pr[k] \simeq \Pr[k+1]$. On the other hand, when ρ is not much larger than n, the end effect results in large $\Pr[k]$ for a small k. Note that the situations of small ρ are often observed in the existing system [2], and the end effect cannot be ignored. Suppose that checkpointing is performed for every n phone calls (i.e., the checkpointing interval is n). Define $\theta_{k,n}$ as the probability that the retrieval is less than k-call outstanding, and N_n as the number of checkpointing operations performed during T. In cellular network engineering, maximizing $\theta_{k,n}$ and minimizing N_n are two conflict goals.

It is easy to derive that

$$\theta_{k,n} = \left[1 - \left(\frac{\lambda}{\lambda+\eta}\right)^{k+1}\right]\left[1 - \left(\frac{\lambda}{\lambda+\eta}\right)^{n}\right]^{-1}$$

for $m = 1$.

Let $\beta[i]$ be the probability that there are i checkpointing operations during T. Then

$$\beta[i] \quad = \quad \sum_{k=0}^{n-1} \int_{t=0}^{\infty} \frac{(\lambda t)^{in+k}}{(in+k)!} e^{-\lambda t} f(t) dt$$

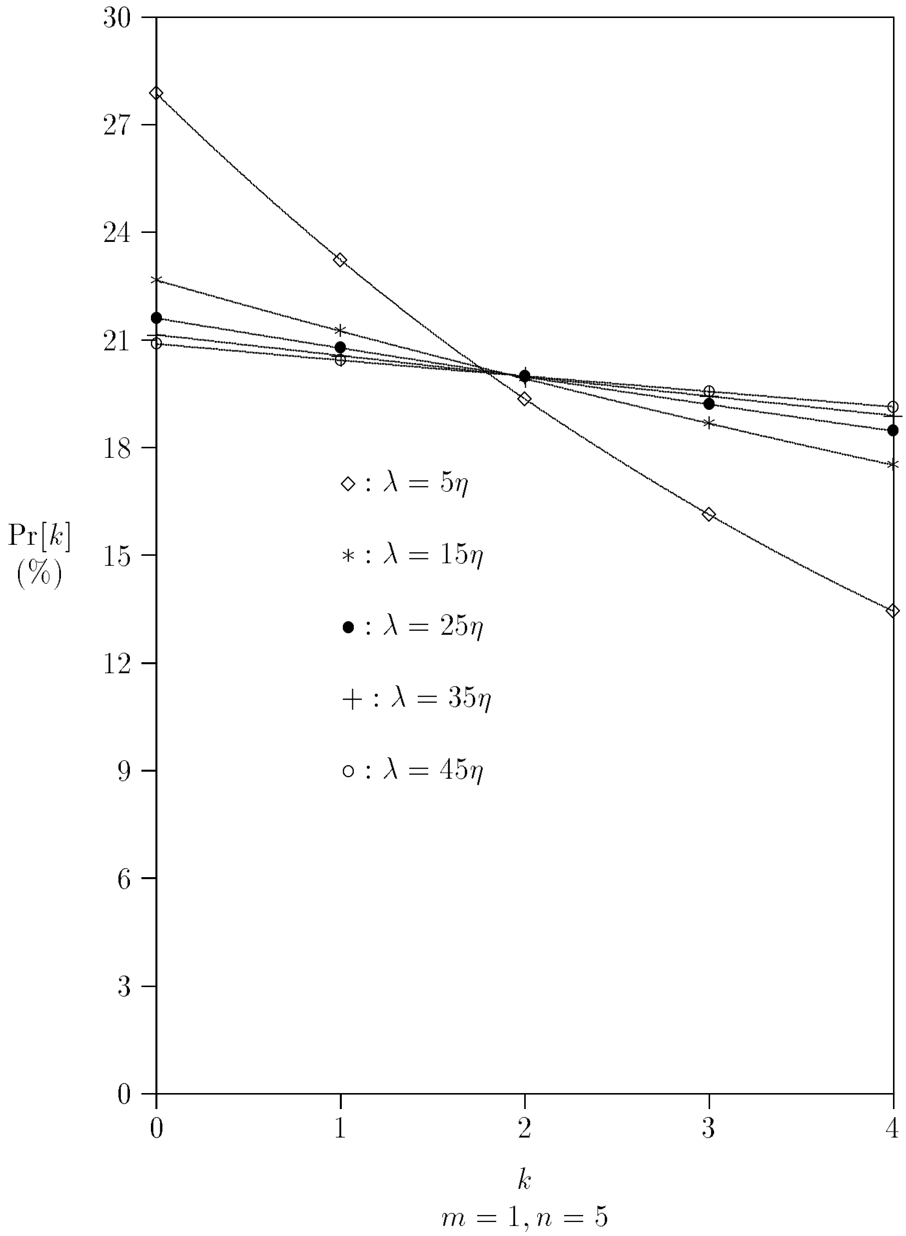


Figure 5.4: Effects of λ on Pr[k] ($n = 5$)

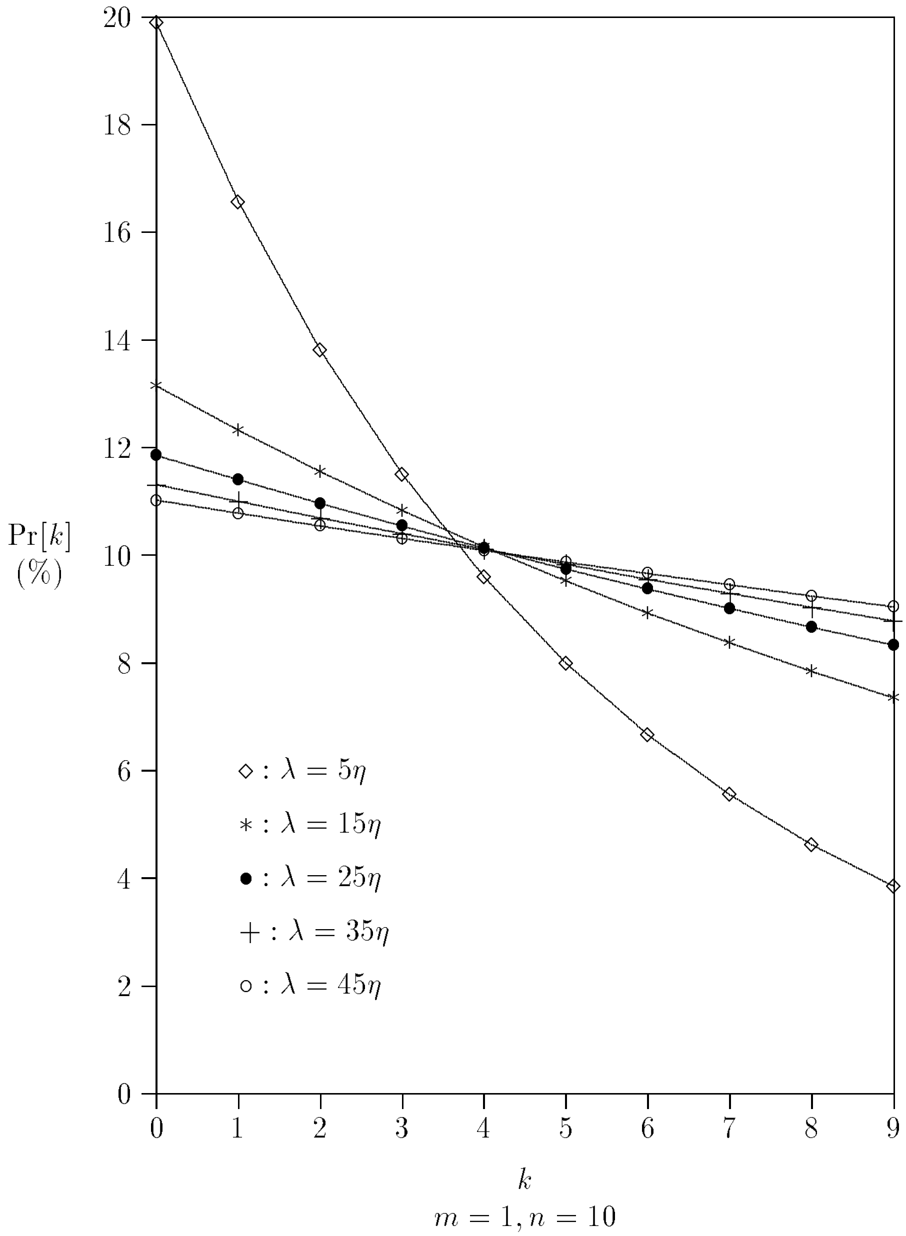


Figure 5.5: Effects of λ on $\Pr[k]$ ($n = 10$)

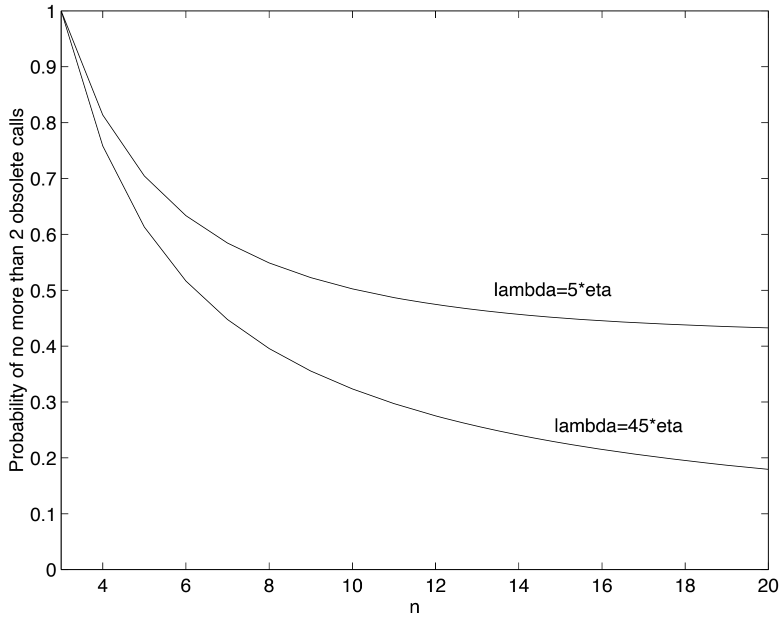


Figure 5.6: The effect of n on $\theta_{2,n}$ $(m = 1)$

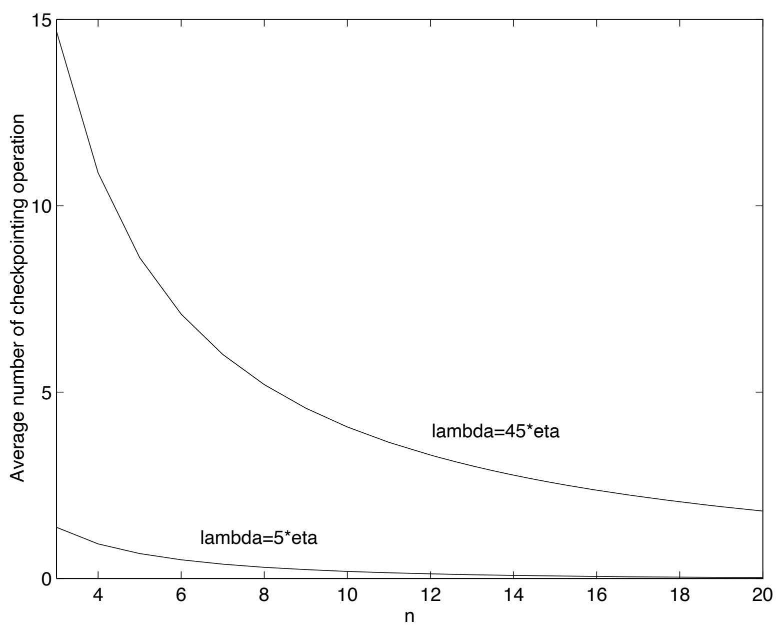


Figure 5.7: The effect of n on $E[N_n]$ $(m = 1)$

$$
\begin{aligned}
&= \sum_{k=0}^{n-1} \frac{\lambda^{in+k}}{(in+k)!} \int_{t=0}^{\infty} t^{in+k} f(t) e^{-\lambda t} dt \\
&= \sum_{k=0}^{n-1} \frac{\lambda^{in+k}}{(in+k)!} (-1)^{in+k} \left[\frac{d^{in+k} f^*(s)}{ds^{in+k}} \right] \Bigg|_{s=\lambda}
\end{aligned}
$$

If $f(t)$ is exponentially distributed (i.e., $m = 1$), then we have

$$
\begin{aligned}
\beta[i] &= \sum_{k=0}^{n-1} \frac{\eta\lambda^{in+k}}{(in+k)!} \frac{(in+k)!}{(\lambda+\eta)^{in+k+1}} \\
&= \frac{\eta\lambda^{in}}{(\lambda+\eta)^{in+1}} \left[\sum_{k=0}^{n-1} \left(\frac{\lambda}{\lambda+\eta} \right)^k \right] \\
&= \left[1 - \left(\frac{\lambda}{\lambda+\eta} \right)^n \right] \left[\left(\frac{\lambda}{\lambda+\eta} \right)^n \right]^i
\end{aligned}
$$

and

$$
\begin{aligned}
E[N_n] &= \sum_{i=0}^{\infty} i\beta[i] \\
&= \left[1 - \left(\frac{\lambda}{\lambda+\eta} \right)^n \right] \sum_{i=0}^{\infty} i \left[\left(\frac{\lambda}{\lambda+\eta} \right)^n \right]^i \\
&= \left(\frac{\lambda}{\lambda+\eta} \right)^n \left[1 - \left(\frac{\lambda}{\lambda+\eta} \right)^n \right]^{-1} \qquad (5.29)
\end{aligned}
$$

Figures 5.6 and 5.7 plot $\theta_{2,n}$ and $E[N_n]$ for $\rho = 5$ and 45, respectively. From these two figures, several engineering questions can be answered. For example, when $\rho = 5$, if n is changed from 12 to 8, then $\theta_{2,n}$ is increased by 16%, and $E[N_n]$ is increased by 130%. With appropriate weighting factors specific to the network under study, one can use the above data to determine whether it is beneficial to change n from 12 to 8. Note that $\theta_{2,n}$ is more sensitive to the change of n when ρ is large than when ρ is small. On the other hand, $E[N_n]$ is more sensitive to the change of n when ρ is small than when ρ is large. When $\rho = 45$, if n is changed from 12 to 8, then $\theta_{2,n}$ is increased by 44%, and $E[N_n]$ is increased by 57%. Thus, it more cost effective to decrease n when ρ is large than when ρ is small.

Next, we consider the billing model based on time interval. Figure 5.8 illustrates the density function for different the average VS residence time $1/\eta$ for the case

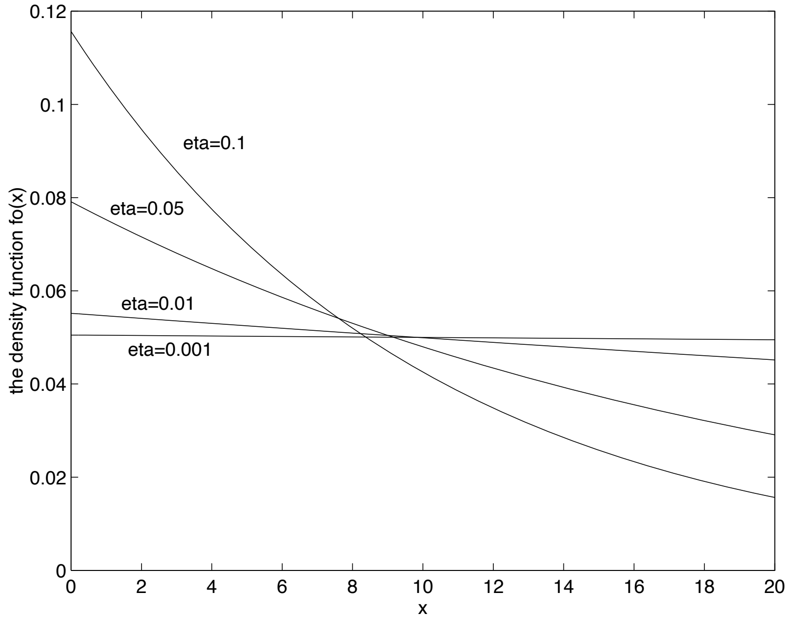


Figure 5.8: The density function of outstanding time for different $\eta (m = 1)$

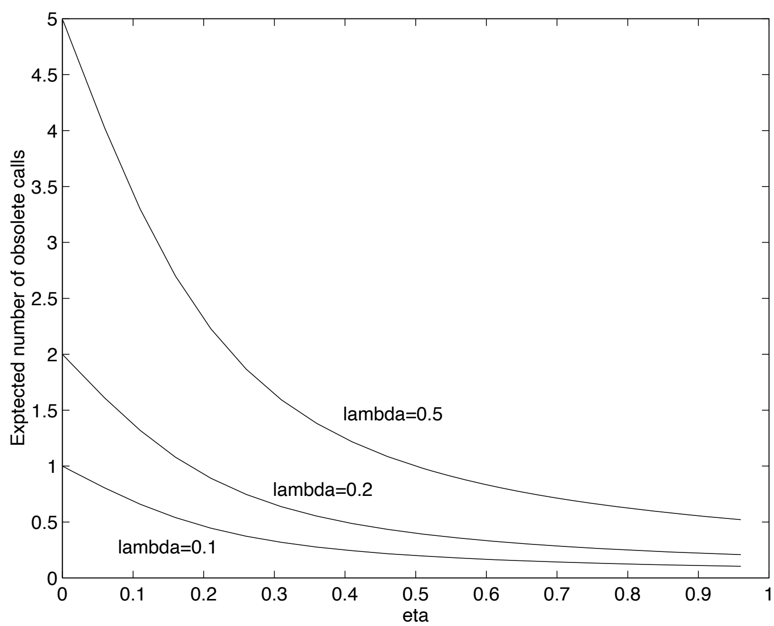


Figure 5.9: The expected number of outstanding calls for different arrival rate $(m = 1)$

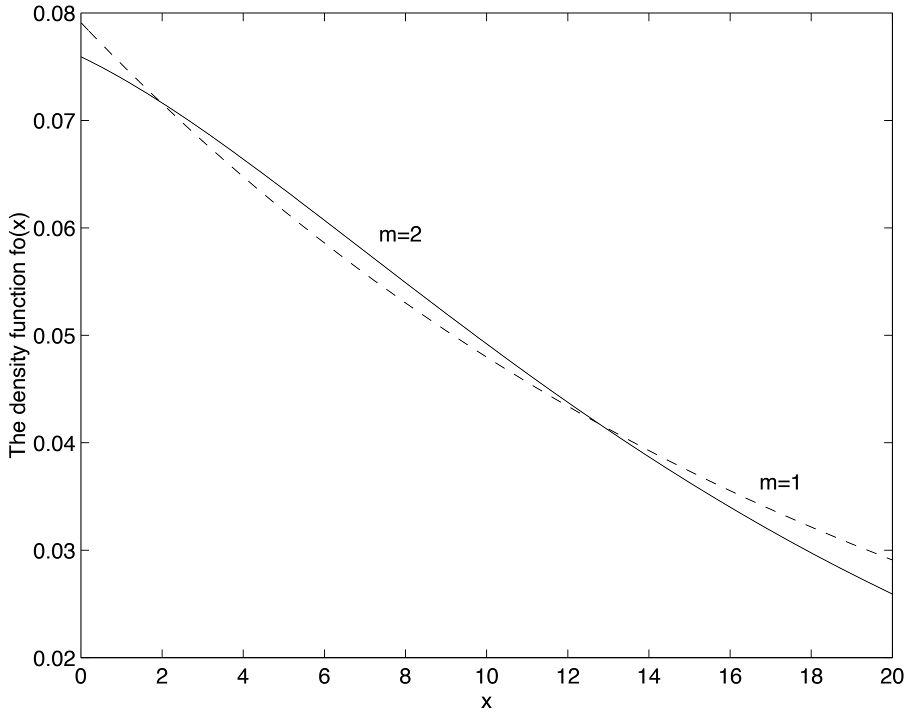


Figure 5.10: The density function for different m

when the VS residence time is exponentially distributed. As in the previous case, the density function is is a decreasing function of time elapsed from the last updating. When the roamer stays sufficiently long at that visited area, the outstanding time is uniformly distributed. When the roamer stays sufficiently short at the visited area, the outstanding time is approximately exponentially distributed.

Figure 5.9 shows the expected number of outstanding calls for different call arrival rate λ. As expected, the expected number of outstanding calls is decreasing as η is increasing, which implies that the shorter the roamer stays, the fewer the outstanding calls. The expected number of outstanding calls is also decreasing as the call arrival rate λ is decreasing, i.e., the fewer the arrival calls, the fewer the outstanding calls. These observations are consistent with our intuition.

Figure 5.10 shows the density function of outstanding time is not much affected by the variance of the VS residence time (which is uniquely determined by m).

5.4 Conclusions

This chapter has proposed a number of strategies for expediting billing record updates in the PCS networks. These strategies have been designed as a compromise between the requirements of traffic for network signaling and the timeliness of the roamer's billing record (i.e., the accuracy of the billing information) at the home system. Performance analysis was performed for all proposed strategies, and was shown to be useful in optimizing the strategies performance by proper setting of timing thresholds to balance control traffic overhead and the usefulness of the billing process. We therefore believe that the results presented here can be used in the future design of billing systems for PCS networks.

Chapter 6

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

PCS networks is a fast-growing area in telecommunications and continue to expand in the next few years. As new technologies, PCS networks have already penetrated telecommunications markets nationally and internationally ([44]). As new demands for integrated services over the air, the PCS networks become more and more integrated and more complex, the system design and analysis become more complicated. New analytical tools, if possible, need to be found. Current trend for performance analysis in wireline and wireless networks is heavily relying on simulation modeling. While this may be helpful for understanding of network performance, however, the end results are usually problem dependent. It will be desirable to provide some analytical tools so that general insights can be drawn.

This research is the first step toward this goal. We have developed new analytical approach for the modeling of the emerging PCS networks. As we have shown, our approach can tackle problems that the classical approach like queueing theory can not solve analytically. This new approach deserves further study and new analytical tools need to be developed along this line. In this section, we list some future research directions.

1. *Heterogeneous PCS Networks*

The basic assumption for some of our results is that the PCS networks are homogeneous, i.e., the mobile customers are uniformly distributed in the service area, hence we use the assumption that the cell residence times are independent and identically distributed. Of course, this is already general enough to approximate the real situation, it may be more desirable to generalize our approach to the heterogeneous PCS networks. This may be too difficult to find a closed-form solution, we may just want to use a less general assumption: the cell residence times are independent but not necessarily identically distributed. For this case, even computational results are helpful.

2. *Cell Traffic and Channel Occupancy Times*

When we study cell traffic and channel occupancy times, we assume that the call holding time is exponentially distributed, where the memoryless property plays a very important role. It will be a very challenging problem to study the cell traffic and channel occupancy times when the call holding time has a more general distribution. We conjecture that the necessary and sufficient condition for the validity of the two commonly used assumptions is that the cell residence time is exponentially distributed, which is true for the case when call holding time is exponentially distributed. Up to now, we have not been able to solve this problem yet.

3. *Blocking Probabilities*

Even for homogeneous PCS networks, we have not been able to find a method to compute the new call blocking probability and handoff call probability under general call holding time and general cell residence times. The loss network theory ([11, 31]) may be used, however, the computation complexity is a major issue considering so huge number of cells used in a PCS network. It seems that decomposition techniques in system theory and perturbation analysis in queueing theory may be helpful at least computationally. Of course, it will be much more desirable to have some analytical results.

4. *Service Charging Planning*

Although we only briefly discuss the service charging planning in Chapter 4, there are plenty of works to be done. Instead of using one parameter—the expected effective call holding time T, we propose planning using two parameters—the expected effective call holding time for a complete call T_c and the expected effective call holding time for an incomplete call T_i. This new approach takes customer care into consideration in mathematical terms. More works need to be done in this area with the cooperation of service providers who are willing to provide their charging schemes and data for this study.

5. *Billing Systems with Fraud Detection*

Roaming services make the PCS billing systems complicated and make wireless fraudulent usage easier ([17]). In order to stay competitive, a PCS service provider has to take customer care into consideration. While appropriate rating schemes may help the cause, a good billing system with fraud detection is another approach to enhance the customer care program. We proposed three billing record updating schemes for billing system monitoring, these schemes can also be modified for possible fraud detection. Performance analysis for two of the schemes were provided. The analysis for the combined scheme is very difficult. It will be very helpful to give an analysis either analytically or by simulations.

6. *Simulation Package and Testbed*

The proposed models for PCS networks in this research can be easily used to build up simulation models. In this package, various distributions for call holding times and cell residence times can be used to evaluate the performance of any type of PCS networks, the performance metrics proposed in this research can be obtained under various traffic conditions. Because of the specific cellular structures of PCS networks, the object-oriented programming languages such as C++ can be easily used to develop this package. Testbed is another practical approach to study the performance of PCS networks, though the size of testbed may be a problem.

7. *Other Applications of This New Approach*

The analytical tools for PCS network performance analysis developed in this research can be used for other problems. As we have already witnessed in the performance of billing strategies, the technique has been used again. Recently, we apply this approach to propose a new location model for location tracking in PCS networks ([16]). Other applications of this approach is worth exploring.

Bibliography

[1] A. Acampora and M. Naghshineh. Control and quality-of-service provisioning in high-speed microcellular networks. *IEEE Personal Comm.* **1**(2), 36-43, Second quarter, 1994.

[2] E. Bond. IS-124 Standard for cellular roaming moving forward with implementation. *Billing World*. pp. 30-37, March/April, 1996.

[3] Bellcore. *Private Communications*. 1995.

[4] V. A. Bolotin. Modeling call holding time distributions for CCS network design and performance analysis. *IEEE J. on Selected Areas in Comm.* **12**(3), 433-438, 1994.

[5] P. J. Burke. The output of a queueing system. *Operations Research*. **4**, pp. 699-704, 1956.

[6] D. C. Cox. Wireless personal communications: what is it? *IEEE Personal Comm. Mag.*, 20-35, April, 1995.

[7] D. R. Cox. *Renewal Theory*. John Wiley & Sons, New York, 1962.

[8] E. Del Re, R. Fantacci and G. Giambene. Handover and dynamic channel allocation techniques in mobile cellular networks. *IEEE Trans. Veh. Tech.* **44**(2), pp. 229-237, 1995.

[9] ETSI. *Digital European Telecommunications Services and Facilities Requirements Specification.* Technical Report ETSI DI/RES 3002, European Telecommunications Standards Institute. 1991.

[10] ETSI/TC, *Mobile Application Part (MAP) Specification*, Version 4.8.0, Recommendation GSM 09.02, 1994.

[11] D. E. Everitt. Traffic engineering of the radio interface for cellular mobile networks. *Proceedings of the IEEE.* **82**(9), 1371-1382, 1994.

[12] Y. Fang, I. Chlamtac and Y. B. Lin. Call performance of a PCS network. To appear in *IEEE Journal on Selected Area in Communications.*

[13] Y. Fang, I. Chlamtac and Y. B. Lin. Modeling PCS networks under general call holding time and cell residence time distributions. Submitted to *IEEE Transactions on Networking.*

[14] Y. Fang, I. Chlamtac and Y. B. Lin. Channel occupancy times and handoff rate for mobile computing and PCS networks. Submitted to *IEEE Transactions on Computers.*

[15] Y. Fang, I. Chlamtac and Y. B. Lin. Billing strategies and performance analysis for PCS networks. Submitted to *IEEE Transactions on Vehicular Technology.*

[16] Y. Fang, I. Chlamtac and Y. B. Lin. Portable movement modeling for PCS networks. In preparation.

[17] E. Finegold. Fighting cellular fraud. *Billing World.* 36-40, March, 1997.

[18] G. J. Foschini, B. Gopinath and Z. Miljanic. Channel cost of mobility. *IEEE Trans. Veh. Tech.* **42**(4), pp. 414-424, 1993.

[19] V. K. Garg and J. E. Wilkes. *Wireless and Personal Communications Systems.* Prentice-Hall PTR, Upper Saddle River, New Jersey, 1996.

[20] R. A. Guerin. Channel occupancy time distribution in a cellular radio system. *IEEE Trans. Veh. Tech.* **35**(3), pp. 89-99, 1987.

[21] D. Grillo. Personal communications and traffic engineering in ITU-T: the developing E.750-series of recommendations. *IEEE Comm. Mag.* 16-28, December, 1996.

[22] N. A. J. Hastings and J. B. Peacock. *Statistical Distributions.* John Wiley & Sons, Inc. 1975.

[23] J. M. Holtzman and D. J. Goodman (ed.). *Wireless Communications: Future Directions.* Kluwer Academic Publishers, Boston, 1993.

[24] D. Hong and S. S. Rappaport. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. *IEEE Trans. Veh. Tech.* **35** (3), 77-92, 1986.

[25] B. Jabbari. Teletraffic aspects of evolving and next-generation wireless communication networks. *IEEE Comm. Mag.* 4-9, December, 1996.

[26] B. Jabbari. Network issues for wireless communications. *IEEE Comm. Mag.* 88-98, January, 1995.

[27] N. L. Johnson. *Continuous Univariate Distributions*: Vol. 1. John Wiley & Sons, Inc. New York. 1970.

[28] R. H. Katz. Improving communication through "situation awareness" adaptation and mobility in wireless information systems. *IEEE Personal Comm.* **1**(1), 6-17, First quarter, 1994.

[29] I. Katzela and M. Naghshineh, Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey. *IEEE Personal Communications.* **3**(3), 10-31, June, 1996.

[30] F. P. Kelly, *Reversibility and Stochastic Networks.* John Wiley & Sons, Inc. New York. 1979.

[31] F. P. Kelly. Loss networks. *The Annals of Applied Probability.* **1**(3), pp. 319-378, 1991.

[32] L. Kleinrock, *Queueing Systems: Theory, Volume I.* John Wiley & Sons, New York, 1975.

[33] D. Lam, D. C. Cox and J. Widom. Teletraffic modeling for personal communications services. *IEEE Comm. Mag.* 79-87, February, 1997.

[34] W. C. Y. Lee, *Mobile Cellular Telecommunications: Analog and Digital Systems*, 2nd Edition, McGraw-Hill, Inc., New York, 1995.

[35] W. R. LePage, *Complex Variables and the Laplace Transform for Engineers.* Dover Publications, Inc., New York, 1980.

[36] Y. B. Lin, L. F. Chang and A. R. Noerpel. A personal communication network handoff model and its applications. Submitted for Publication. 1996. Also available at *http://liny.csie.nctu.edu.tw/*.

[37] Y. B. Lin and S. K. DeVries. PCS network signaling using SS7. *IEEE Personal Communications Magazine.* pp. 44-55, June, 1995.

[38] Y. B. Lin and I. Chlamtac. Heterogeneous personal communications services. *IEEE Communications Magazine*, **34**(9), pp. 106-113, September, 1996.

[39] Y. B. Lin and I. Chlamtac. Effective call holding times for a PCS network. Submitted to *IEEE JSAC.* 1996.

[40] Y. B. Lin, S. Mohan and A. Noerpel. Queueing priority channel assignment strategies for handoff and initial access for a PCS network. *IEEE Trans. Veh. Technol.*, **43**(3), 704-712, 1994.

[41] Y. B. Lin. Signaling system number 7. To appear in *IEEE Potentials*. See also http://liny.csie.nctu.edu.tw.

[42] Y. B. Lin. Reducing location update cost in a PCS network. To appear in *IEEE/ACM Transactions on Networking*, 1996. Also available at *http://liny.csie.nctu.edu.tw/*.

[43] Y. B. Lin, A. Noerpel and D. Harasty. The sub-rating channel assignment strategy for PCS Hand-offs. *IEEE Trans. Veh. Technol.*, **45**(1), Feb., 1996.

[44] J. Lucas. The top 10: winners and losers in the new wireless marketplace. *Billing World*. March, 1997.

[45] F. Mademann. General packet radio services. *Telecom Report*. **18**, pp. 42-45, 1995.

[46] J. Medhi. *Stochastic Models in Queueing Theory*. Academic Press, New York. 1991.

[47] I. Mitrani. *Modeling of Computer and Communication Systems*. Cambridge University Press, 1987.

[48] S. Mohan and R. Jain. Two user location strategies for personal communications services. *IEEE Personal Communications*, 42-50, First quarter, 1994.

[49] S. H. Moran. Cellular companies to expense up to $1.2 billion on billing and customer care in 1996. *Billing World*. pp. 12-16, March/April, 1996.

[50] S. Nanda. Teletraffic models for urban and suburban microcells: cell sizes and handoff rates. *IEEE Trans. Veh. tech.* **42**(4), 673-682, 1993.

[51] A. R. Noerpel, L. F. Chang and Y. B. Lin. Polling deregistration for unlicensed PCS. *IEEE Journal on Selected Area in Communications*. **14**(4), pp. 728-734, 1996.

[52] A. R. Noerpel, Y. B. Lin and H. Sherry. PACS: Personal access communications system–a tutorial. *IEEE Personal Communications.* **3**(3), 32-43, June, 1996.

[53] J. E. Padgett, C. G. Gunther and T. Hattori. Overview of wireless personal communications. *IEEE Comm. Mag.*, 28-41, January, 1995.

[54] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*, John Wiley & Sons, Inc., New York, 1995.

[55] R. Pandya. Emerging mobile and personal communication systems. *IEEE Comm. Mag.* 44-52, June, 1995.

[56] G. P. Pollini. Trends in handover design. *IEEE Comm. Mag.*, 82-90, March, 1996.

[57] Race Industrial Consortium. *Race-UMTS requirements specs.* Doc.No.16/91, ETSI-SMG5, Oslo, Norway, November, 1991.

[58] S. S. Rappaport and L. Hu. Microcellular communication systems with hierachical macrocell overlays: traffic performance models and analysis. *Proceedings of the IEEE.* **82**(9), 1383-1397, 1994.

[59] S. S. Rappaport. The multiple-call handoff problem in high-capacity cellular communications systems. *IEEE Trans. Veh. Tech.* **40**(3), 546-557, 1991.

[60] S. S. Rappaport. Blocking, hand-off and traffic performance for cellular communication systems with mixed platforms. *IEE Proceedings–I*, **140**(5), pp. 389-401, 1993.

[61] S. S. Rappaport and C. Purzynski. Prioritized resource assignment for mobile cellular communication systems with mixed services and platform types. *IEEE Transactions on Vehicular Technology.* **45**(3), pp. 443-458, 1996.

[62] J. Riordan. *Combinatorial Identities.* John Wiley & Sons, 1968.

[63] S. M. Ross. *Introduction to Probability Models*. Fifth Edition. Academic Press, New York, 1993.

[64] R. Steedman. The common air interface MPT 1375. *Cordless Telecommunications in Europe* (W. H. W. Tuttlebee ed.), Springer-Verlag, New York. 1990.

[65] R. Steele, J. Whitehead and W. C. Wong. System aspects of cellular radio. *IEEE Comm. Mag.* 80-86, January, 1995.

[66] S. Tekinay and B. Jabbari. A measurement-based prioritization scheme for handovers in mobile cellular networks. *IEEE J. Selected Area in Comm.* **10**(8), 1343-1350, 1992.

[67] R. Vijayan and J. M. Holtzman. A model for analyzing handoff algorithms. *IEEE Trans. on Vehi. Tech.* **42**(3), 351-356, 1993.

[68] W. C. Wong. Packet reservation multiple access in a metropolitan microcellular radio environment. *IEEE J. Selected Area in Comm.* **11**(6), 918-925, 1993.

[69] H. Xie and S. Kuek. Priority handoff analysis. *Proc. of Vehicular Technology Conference*. 855-858, 1993.

[70] K. L. Yeung and S. Nanda. Channel management in microcell/macrocell cellular radio systems. *IEEE Trans. Veh. Tech.* **45**(4), 601-612, 1996.

[71] C. H. Yoon and C. K. Un. Performance of personal portable radio telephone systems with and without guard channels. *IEEE J. Selected Areas in Comm.* **11**(6), pp. 911-917, 1993.

[72] T-S. P. Yum and K. L. Yeung. Blocking and handoff performance analysis of directed retry in cellular mobile systems. *IEEE Trans. Veh. Tech.* **44**(3), pp. 645-650, 1995.

Appendix A

Curriculum Vitae

Yuguang Fang, Ph.D

38 Westwind Road #310

Dorchester, MA 02125

Tel: (617)265-9715

Email: yxf2@bu.edu

Degrees

Ph.D in Electrical and Computer Engineering, majoring in Wireless Networks and Mobile Communications, **Boston University**, Boston, Massachusetts, May, 1997.

- *Ph.D Dissertation*: "Personal Communications Services (PCS) Networks: Modeling and Performance Analysis."

Ph.D in Systems and Control Engineering, **Case Western Reserve University**, Cleveland, Ohio, Jan., 1994.

- *Ph.D Dissertation*: "Stability Analysis of Linear Control Systems with Uncertain Parameters."

M.S. in Mathematics, **Qufu Normal University**, Shandong, PRC, 1987.

B.S. in Mathematics, **Qufu Normal University**, Shandong, PRC, 1984.

Experience

September, 1995–May, 1997, **Research Assistant, Department of Electrical and Computer Engineering, Boston University.** Working on Wireless and Mobile Communications Networks, including cellular and PCS networks.

June, 1994–September, 1995, **Postdoctoral Research Associate, Department of Electrical and Computer Engineering, Boston University.** Research on Neurodynamics: Modeling, Analysis and Synthesis of Neural Networks with applications to optimization and neural circuit design.

Jan. 1994–May, 1994, **Research Associate, Department of Systems and Control Engineering, Case Western Reserve University, Cleveland, Ohio.** Working on Fault Detection Analysis of Rotor Dynamics.

Aug. 1989–Jan. 1994, **Teaching and Research assistant, Case Western Reserve University**. Research areas and projects include: Fault Detection Analysis of Rotor Dynamics; Geometrical Approach to Nonlinear Systems and Variable Structure Systems; Robust Stability and Control; Stability of Stochastic Systems, Feedback Control Systems with Communication Delay, Modeling and Analysis of Control Systems in Communication Networks (CASE/FORD projects) supported by the Ford Motor Company; Neural Networks Theory and Applications in Drug Delivery Systems; Drug Delivery Systems and Adaptive Predictive Control.

July, 1987–Oct., 1988, **Teaching and Research Faculty, Department of Mathematics & The Institute of Automation, Qufu Normal University, Shandong, PRC.** Teaching Courses: Model Reference

Adaptive Control; Adaptive Filtering, Prediction and Control; Participating in Research Projects related to Medical Systems.

Professional Activities

Member of the IEEE.

Active reviewers for various professional journals and conferences.

Publications

Most Recent Research Papers

1. Y. Fang, I. Chlamtac and Y. B. Lin, "Call Performance for a PCS Networks," To appear in *IEEE Journal on Selected Area in Communications.*
2. Y. Fang, I. Chlamtac and Y. B. Lin, "Modeling PCS Networks Under General Call Holding Times and Cell Residence Time Distributions," Submitted to *IEEE Transactions on Networking.*
3. Y. Fang, I. Chlamtac and Y. B. Lin, "Channel Occupancy Times and Handoff Rate for Mobile Computing and PCS Networks" Submitted to *IEEE Transactions on Computers.*
4. Y. Fang, I. Chlamtac and Y. B. Lin, "Billing Strategies and Performance Analysis for PCS Networks," Submitted to *IEEE Transactions on Vehicular Technology.*
5. Y. Fang, I. Chlamtac and Y. B. Lin, "Portable movement modeling for PCS Networks," Submitted to *IEEE Transactions on Communications.*

Journal Publications

1. Y. Fang, "A General Sufficient Condition for Almost Sure Stability of Jump Linear Systems," *IEEE Transactions on Automatic Control*, Vol. 42, No. 3, pp. 378-382, 1997.

2. Y. Fang and T. G. Kincaid, "Global Properties for a Class of Dynamical Neural Circuits," Accepted for Publication in *Journal of the Franklin Institute*.

3. Y. Fang and T. G. Kincaid, "Stability Analysis of Dynamical Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 7, No. 4, pp. 996-1006, July, 1996.

4. Y. Fang, M. Cohen and T. G. Kincaid, "Dynamics of a Winner-Take-All Dynamical Neural Network," *Neural Networks*, Vol. 9, No. 7, pp. 1141-1154, 1996.

5. Y. Fang, K. A. Loparo and X. Feng, "Stability of Discrete Time Jump Linear Systems," *Journal of Mathematical Systems, Estimation and Control*, Vol. 5, No. 3, pp. 275-321, July, 1995.

6. Y. Fang, "The Analysis of Eigenvalue Assignment Robustness," *IEEE Transactions Automatic Control*, Vol. 40, No. 6, pp. 1108-1109, June, 1995.

7. X. Feng, K. A. Loparo and Y. Fang, "Optimal State Estimation with Active Probing for Stochastic Systems: An Information Theoretic Approach," Accepted for Publication in *IEEE Transaction on Automatic Control*.

8. Y. Fang, K. A. Loparo and X. Feng, "Inequalities for the Trace of Matrix Product," *IEEE Transactions on Automatic Control*, Vol. 39, No. 12, pp. 2489-2490, 1994.

9. Y. Fang, K. A. Loparo, X. Feng, "Almost Sure and $\delta-$moment Stability of Jump Linear Systems," *International Journal of Control*, Vol. 59, No. 5, pp. 1281-1307, 1994.

10. Y. Fang, K. A. Loparo and X. Feng, "A Sufficient Condition for Stability of a Polytope of Interval Matrices," *Systems and Control Letters,* Vol. 23, No. 4, pp. 237-245, 1994.

11. Y. Fang, K. A. Loparo and X. Feng, "Sufficient Conditions for the Stability of Interval Matrices," *International Journal of Control,* Vol. 58, No. 4, pp. 969-977, 1993.

12. Y. Fang, K. A. Loparo and X. Feng, "New Estimation Results for the Solution of Lyapunov Equations," *IEEE Transactions on Automatic Control*, Vol. 42, No. 3, pp. 408-411, 1997.

Submitted Papers for Journal Publications

1. Y. Fang, M. Cohen and T. G. Kincaid, "A New Class of Winner-Take-All Competitive Dynamical Neural Networks," Submitted to *IEEE Transactions on Neural Networks.*

2. Y. Fang and T. G. Kincaid, "Sufficient Conditions for Stability of a Class of Dynamical Neural Networks," Submitted to *IEEE Transactions on Neural Networks.*

3. Y. Fang, "Generalization of Mapping Theorem and its Applications to Robust Stability Analysis," Submitted to *IEEE Transactions on Automatic Control.*

4. Y. Fang, "New Robust Stability Results for Linear Systems with Structured Perturbation," Submitted to *Systems and Control Letters.*

5. Y. Fang, "New Robust Stability Criteria for Linear Systems with Multiple Time Delays and Nonlinear Perturbation," Submitted to *IEEE Transactions on Automatic Control.*

6. Y. Fang, "Robust Stability of Uncertain Linear Systems with Time Delay," Submitted to *Systems and Control Letters.*

7. Y. Fang, "Robust Stability Analysis for a Class of Uncertain Nonlinear Systems," Submitted to *International Journal of Control.*

8. Y. Fang, "Robust Stability of Linear Systems with Normed-bounded Perturbation," Submitted to *IEEE Transactions on Automatic Control*.

9. Y. Fang, K. A. Loparo and X. Feng, "Robust Stability and Performance Analysis: A New Approach," Submitted to *Systems and Control Letters*.

10. Y. Fang and K. A. Loparo, "Stability Robustness Bounds and Robust Stability for Linear Systems with Structured Uncertainty," Submitted to *IEEE Transactions on Automatic Control*.

11. Y. Fang and T. G. Kincaid, "Convergence Analysis of the Widrow-Hoff Algorithm for McCulloch-Pitts Neurons with Nondecreasing Activations," submitted to *IEEE Trans. Neural Networks*.

12. Y. Fang, K. A. Loparo and X. Feng, "Stability of Continuous-time Jump Linear Systems," submitted to *Stochastics and Stochastic Reports*.

13. Y. Fang, K. A. Loparo and X. Feng, "Stochastic Stability of Jump Linear Systems," submitted to *IEEE Transaction on Automatic Control*.

14. Y. Fang, K. A. Loparo and X. Feng, "Sufficient Conditions for Robust Stability of a Polytope of Discrete Systems: A Unified Approach," submitted to *International Journal of Control*.

15. Y. Fang, K. A. Loparo and X. Feng, "On Relations between the Sample and Moment Lyapunov Exponents of Stochastic Systems," Submitted to *Stochastics and Stochastic Reports*.

16. Y. Fang, K. A. Loparo and X. Feng, "Stabilization of Continuous-time Jump Linear Systems," Submitted to *IEEE Transactions on Automatic Control*.

17. Y. Fang and K. A. Loparo, "Exponential Stabilization and Exponential Observers of Nonlinear Systems," Submitted to *Control– Theory and its Applications*.

18. Y. Fang and K. A. Loparo, "Robust Stability Bounds of Discrete-time Linear Systems with Structured Perturbations," Submitted to *IEEE Transactions on Automatic Control*.
19. Y. Fang, K. A. Loparo and X. Feng, "Robust Stability Analysis of Uncertain Control Systems via Matrix Measure," Submitted to *International Journal of Control*.

Conference Papers

1. Y. Fang and T. G. Kincaid, "Sufficient Conditions for Stability of a Class of Dynamical Neural Circuits," *1996 Triennial World Congress of the International Federation of Automatic Control (IFAC'96)*, San Francisco, California, USA, June 30-July 5, 1996.
2. Y. Fang, "Almost Sure Stability of Jump Linear Systems," *1996 Triennial World Congress of the International Federation of Automatic Control (IFAC'96)*, San Francisco, California, USA, June 30-July 5, 1996.
3. Y. Fang, T. G. Kincaid and M. Cohen, "Three Classes of Winner-Take-All Neural Networks," *Proceedings of World Congress on Neural Networks (WCNN'95)*, Washington, D. C., Vol. I, pp. 266-269, July 17-21, 1995.
4. Y. Fang and T. G. Kincaid, "Global Properties for a Class of Dynamical Neural Circuits," *Proceedings of the American Control Conference*, Seattle, pp. 2762-2766, 1995.
5. Y. Fang and K. A. Loparo, "Stability Robustness Bounds and Robust Stability for Linear Systems with Structured Uncertainty," *Proceedings of the American Control Conference,* Baltimore, Maryland, pp. 221-225, June, 1994.
6. Y. Fang and X. Chu, "Stability Analysis of Multirate Sampled-data Systems," *Proceedings of the American Control Conference*, Baltimore, Maryland, June, 1994.

7. Y. Fang, K. A. Loparo and X. Feng, "Robust Stability and Performance Analysis: A New Approach," *Proceedings of the 32th IEEE Conference on Decision and Control,* San Antonio, Vol. 3, pp. 2006-2007, 1993.

8. Y. Fang, K. A. Loparo and X. Feng, "New Estimation Results for the Solution of Lyapunov Equations," *Proceedings of the 32th IEEE Conference on Decision and Control*, San Antonio, Vol. 4, pp. 3208-3209, 1993.

9. X. Chu, G. Du, Y. Fang and H. Peng,"A Closed-loop Self-adaptive Fluid Infusion of Burn Patients," *The Eighth IFAC/IFORS Symposium on System Identification and Parameter Estimation*, Aug. 26-31, 1988, Beijing, Pergamon Press.

Technical Reports

1. Y. Fang, K. A. Loparo and X. Feng, "SCP Queue Modeling," SCP Research Report for Ford (1), Systems Engineering Dept., CWRU, Cleveland, Ohio, July, 1991.

2. Y. Fang, K. A. Loparo and X. Feng, "Modeling Issues for the Control Systems with Communication Delays," SCP Research Report for Ford (2), Systems Engineering Dept., CWRU, Cleveland, Ohio, Oct., 1991.

3. Y. Fang, K. A. Loparo and X. Feng, "SCP Queue Modeling and Almost Sure Stability of Control Systems with Communication Delays," SCP Research Report for Ford (3), Systems Engineering Dept., CWRU, Cleveland, Ohio, Jan., 1992.

4. Y. Fang, "Optimal Selection of the Prediction Horizon for Adaptive Controllers Using Artificial Neural Networks," Technical Research Report, Department of Systems Engineering, Case Western Reserve University, Cleveland, OH 44106, Dec. 1990.

5. Y. Fang, “On Layered Neural Networks with Gaussian Activations,” Technical Research Report, Department of Systems Engineering, Case Western Reserve University, Cleveland, OH 44106, Dec. 1990.