MODELING AND CHARACTERIZATION OF TRAFFIC IN PUBLIC SAFETY WIRELESS NETWORKS

by

Božidar Vujičić
Dipl.Ing.E.E., University of Montenegro, 1997

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

in the
School
of
Engineering Science

SIMON FRASER UNIVERSITY

Fall 2006

APPROVAL

Name: **Božidar Vujičić**

Degree: **Master of Applied Science**

Title of Thesis: **Modeling and characterization of traffic in public safety wireless networks**

Examining Committee:

Chair: **Dr. Jie Liang**
Assistant Professor of School of Engineering Science

____

Dr. Ljiljana Trajkovic
Senior Supervisor
Professor of School of Engineering Science

____

Dr. Uwe Glässer
Supervisor
Associate Professor of School of Computing Science

____

Dr. Stephen Hardy
Internal Examiner
Professor of School of Engineering Science

Date Defended/Approved: Nov. 28, 2006

ABSTRACT

Reliable communication and interoperability between public safety agencies play vital role for public safety. We analyze traffic data from a deployed trunked radio communication network operated by public safety wireless network service provider E-Comm. Traffic data span various periods in 2001, 2002, and 2003. OPNET model is created to evaluate the utilization of network resources and to locate network bottlenecks. Our analysis may be used to address existing and future network congestion problems. We also present statistical analysis of traffic data. We examine statistical distribution and autocorrelation function of call inter-arrival and call holding times during several busy hours. We find that call inter-arrival times are long-range dependent and may be modelled by both Weibull and gamma distributions. Call holding times follow the lognormal distribution and are uncorrelated. These findings indicate that traditional Erlang models for voice traffic may not be suitable for evaluating the performance of trunked radio networks.

Keywords: Trunked radio systems, emergency communications, wireless networks, network utilization, traffic analysis, long-range dependence, wavelets.

Subject Terms: Traffic modeling, emergency communications, public safety wireless network.

This thesis is dedicated,

to my mother Jelena, and in loving memory of my father Slobodan

who introduced me to the joy of learning from birth

and taught me the value of education enabling this study to take place today.

This work is result of the devoted care

that good parents provide for their offspring

that opens the way for their children to grow in wisdom

and to see and believe in the possibilities in life.

I hope that this scientific undertaking and

the effort required to complete this degree

will serve as motivation for my son Stefan

for his own future academic and life endeavours,

whatever they may be.

This thesis is also dedicated to my sole mate,

my wife, Lankica, whose unconditional love, sacrifice, patience, guidance, support,

and countless other little things, all of which only her love could endure ,

always were and will be an inspiration for me

At that time, Jesus spoke these words:

"I thank thee, Father, Lord of heaven and earth,

for hiding these things from the learned and wise,

and revealing them to the simple ..."

The Gospel according to Matthew, ch. 11

"No matter what you've done for yourself or for humanity,

if you can't look back

on having given love and attention to your own family,

what have you really accomplished?"
– Lee Iacocca

ACKNOWLEDGEMENTS

"Often we can help each other most by leaving each other alone; at other times we need the hand-grasp and the word of cheer"- Elbert Hubbard.

I very much appreciate "the hand-grasp" and "the word of cheer" from my senior supervisor Dr. Ljiljana Trajkovic and all members of the Communication Networks Laboratory at Simon Fraser University. I would like to use this opportunity to express my gratitude to Dr. Trajkovic for the academic support, continuously guidance, insight, trust and above all, her everlasting patience, warm-heartedness, and emotional support in difficult moments. From her, I learned how to conduct research and how to publish research results. It was a great pleasure for me to work with my colleagues in a productive environment of the Communication Networks Laboratory under her supervision. I express sincere appreciation for the fellow graduate students from our Communication Network Laboratory (CNL) for valuable comments and enlightening discussion and for being such wonderful friends. I especially emphasise my thankfulness to Nikola Cackov for being a great colleague and friend.

I also thank Duncan Sharp from Planetworks and the management and technical staff at E-Comm for providing access to the activity data and technical support for data analysis. Thanks are also due to the anonymous reviewers of my published papers for their valuable comments and suggestions.

I would like to express appreciation to my examining committee: Dr. Uwe Glässer for his constructive suggestions and detailed and perceptive comments, Dr.

Stephen Hardy for inspiring discussions on traffic modelling, and Dr. Jie Liang for chairing the thesis defence.

My special thanks go to my dear professors and friends from University of Montenegro, for all their kindness and confidence in me. Thanks for giving me the foundation to be who I am and for recommendation and advices to pursue this studies. I extend thanks to my friends and colleagues from Trebinje for their years of love and friendship that have come to value.

This endeavour to complete a Masters degree could never have been accomplished without all the love and support of my wife and my family over the years. Although miles away from me, they were always by my side. Their sacrifice, patience and endurance have been a gift that I will admire always.

Finally, I would be remiss without mentioning my aunts Zorka and Danica Vujicic whose extreme generosity and love will be remembered always.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ARIMA	Auto Regressive Integrated Moving Average
CDF	Cumulative Density Function
DARPA	Defense Advanced Research Projects Agency
DWT	Discrete Wavelet Transform
E-Comm	Emergency Communications for Southwestern British Columbia Inc.
ECDF	Empirical Distribution Function
EDACS	Enhanced Digital Access Communications System
FIFO	First In First Out
FSM	Finite State Machine
GoS	Grade of Service
GVRD	Great Vancouver Regional District
K-S	Kolmogorov-Smirnov Goodness-of-Fit Test
LMR	Land Mobile Radio
LRD	Long Range Dependence
MLE	Maximum Likelihood Estimation
PBX	Private Branch Exchange
PDF	Probability Density Function
PAMR	Public Access Mobile Radio
PCS	Personal Communication Services
PMR	Private Mobile Radio System
PSD	Power Spectral Density

PSTN	Public Switched Telephone Network
PSWN	Public Safety Wireless Network
PTT	Push-to-Talk
SDR	Software Defined Radios
WARN	Wide Area Radio Network
WarnSim	Wide Area Radio Network Simulator

CHAPTER 1: INTRODUCTION

Simulation and analysis of traffic in deployed communication networks are used to determine their operational status, their performance, and to identify and locate possible network congestion. Traffic modelling is necessary for network provisioning, predicting utilization of network resources, and for planning network developments. These studies may be used to improve network reliability, which is particularly important for networks used by public safety agencies, such as police, fire department, and ambulance [1]–[4]. The scope of the performance evaluation and the parameters of interest depend on the network and its characteristics, such as technology, topology, and user behaviour [5]–[7].

The public safety community has recognized public safety agencies interoperability and the limited and fragmented radio spectrum as main concerns related to operations of public safety wireless communications. The high cost of deploying and operating voice channels is equipoise to coverage and capacity of the network. In an effort to improve public safety response of local, state, and federal public safety agencies through more effective and efficient interoperable wireless communications, understanding, analyzing, evaluating, and optimizing already deployed public safety wireless networks is a very important research area.

E-Comm, Emergency Communications for Southwest British Columbia Inc. [8], provides voice and data transmission services via a Public Safety Wireless Network

(PSWN) to mobile users who belong to various public safety agencies. The E-Comm network is a circuit-switched macro-cellular network [4]. Each cell covers a certain geographical area within the Greater Vancouver Regional District, the Sunshine Coast Regional District, Whistler, and Pemberton, in British Columbia, Canada. The number of frequencies available in each of the eleven cells is predefined and it determines the cell capacity. It corresponds to the number of available radio channels in each cell. Individual radio users access the network (channels) via trunking [9], [10]. It implies sharing a set of frequencies (radio channels) among the agencies rather than dedicating subsets of frequencies to individual agencies.

Previous results on performance evaluation of cellular trunked radio systems have been obtained by developing mathematical models based on traditional queuing theory [11] and statistical analysis of collected traffic [3], [12]–[36]. Traffic in telephone networks was modelled by M/M/n queues: Poisson arrivals (exponentially distributed inter-arrival times), exponential call holding times, and n parallel servers [37]. Based on this theory, voice traffic in circuit-switched networks has been modelled using the Erlang B and C models [38], [39]. These models assume independent and exponentially distributed call holding and call inter-arrival times. The Erlang B is a more general model, and its Grade of Service (GoS) is insensitive to the distribution of service times. These models have proved appropriate for modelling telephone traffic. Trunked radio systems possess characteristics that distinguish them from the telephone networks, such as trunking-based network access and one-to-many type of conversation. Therefore, the Erlang models may not capture the statistical characteristics of the traffic in trunked radio systems. Our analysis of traffic from the E-Comm network shows that neither call holding nor call inter-arrival times are exponentially distributed [40]. Furthermore, we show that call inter-arrival times exhibit long-range dependence.

A number of statistical models have been developed for call holding time in various networks, such as public switched telephone network (PSTN) and personal communication services (PCS) networks. Few models are available for wide area radio communication networks because they are used mainly by public safety agencies. Furthermore, traffic data from PSWNs are strictly confidential and hardly available. The large difference between average call holding times in PSWNs and in PSTN/Cellular networks suggests that call holding time models developed for other networks might not be suitable for PSWNs [1], [2], [4].

Previous study indicated that the channel holding time for common channel signaling (CSS) networks is not best represented using the exponential distribution. A mixture of some well known distributions provides a better model for call holding time. Furthermore, it is shown that channel throughput drops more significantly under an exponential call holding time distribution model than the measured call holding time distribution. The exponential distribution underestimates the contribution of short calls in traditional PSTNs [13]. In traditional PSTNs, the call holding time fits much better the mixture of lognormal distributions with a mean value of 200-300 s than the widely used exponential distribution [14]. The channel holding time in the analyzed micro-cellular network, where handoffs are frequent, follows the lognormal distribution [15].

There are also numerous research results in the area of wireless and mobile computing that indicate that call holding times and call inter-arrival times of cell traffic are no longer exponentially distributed [15]-[34]. For example, the analysis of call holding times from field data for personal communication services collected in the office buildings and residence areas in Taiwan revealed that the call holding time could not be modeled by exponential distribution. Gamma and lognormal distributions provide the second order approximation to traffic data [17]. The channel occupancy times are exponentially distributed if and only if the cell residence times are exponentially distributed. In mobile wireless networks, the cell residence time is defined as the time interval during which a mobile user resides in a cell. (A cell is the area of radio coverage of one base station). Furthermore, the merged traffic from new calls and handoff calls is Poisson if and only if the cell residence times are exponentially distributed [18]. A new mobility model, called hyper-Erlang distribution model [19]-[21], provides analytically tractable queuing system while still fitting field data. It is convenient for the characterization of systems with mixed types of traffic. The effect of this model on channel holding time was also analyzed. The call holding times were modeled by the Erlang distribution (a generalization of the exponential distribution). The model was used to examine the effect of the variance of call holding times on the call completion probability [22].

The Public Access Mobile Radio (PAMR) system [23]-[30] is a macro-cellular network similar to the E-Comm network. It employs a push-to-talk mechanism for network access and it is used by groups working in public transportation and distribution. The PAMR system operates based on two levels of management: messages or transmissions. The message trunking treats the entire conversation as one call, while transmission trunking treats each transmission as a separate unit. The analysis of message and transmission durations revealed that message length (channel holding time) and transmission length may be modeled by Erlang-jk and lognormal distributions (or a mixture of two lognormal distributions), with mean value of 20-40s, respectively [23]-[28]. These results are relevant to findings reported for traditional PSTNs [14]. Analysis of call holding time distributions [14] was based on two concepts: the human perception of time on a logarithmic scale and the fact that the call holding time in a call mix (partial dialling, subscriber busy, no answer) is a combination of various distributions. These studies were based on empirical data of various call types. Each individual component of the entire circuit holding time may be modeled as a mixture of two or more distributions. Barcelo [29] proved that call arrival process of the merged new calls and handover traffic are smoother than Poisson traffic. An assumption of Poisson call arrival overestimates the system's Grade of Service (GoS) [30].

Orlik and Rappaport [31], [32] modeled the cell residence time as the sum of the hyper-exponential (SOHYP) random variables. They showed that the model of the channel holding time fits a large number of statistical distributions. It was shown that the cell residence time can be approximated by generalized gamma distribution [33]. An analytical model used for the performance evaluation of cellular communication network systems was also proposed [34].

Analysis of channel utilization is a common approach for allocating network resources. Industry Canada channel loading guidelines for land mobile radio systems used by safety services recommend below 50% channel occupancy in conventional systems and a 3% probability that calls will not be delayed more than one call holding time in trunked systems during the average busy period [41].

We adopted a simulation approach to model the utilization of the E-Comm network based on collected traffic data. The basic circuit-switched functionalities and performance of the E-Comm network [8] has been evaluated by using the OPNET network simulator [1], [2], [42]. Simulations of the network utilization during two sample weeks addressed the increase in the network traffic volume and network congestion. In order to simulate events characteristic to circuit-switched PSWNs, such as call queuing and retrying of blocked calls, a customized simulation tool, the Wide Area Radio Network Simulation tool (WarnSim), was developed [4], [43]. Furthermore, clustering of network users based on their activity and a seasonal autoregressive integrated moving average (SARIMA) model for predicting traffic from each cluster were proposed in [35], [36].

The objective of this work is to develop statistical models for call traffic in the E-Comm network, to evaluate performance of the E-Comm network in terms of channel utilization and call blocking probability, and to provide a model for predicting future performance of the E-Comm network using the customized OPNET network simulation tool.

The activity data collected by E-Comm were used to examine network utilization over sample weeks in 2002 and 2003. The E-Comm activity data table contains records of network events. The relevant data were processed into a format suitable for OPNET trace-driven network simulations. We examined the instantaneous utilization of radio channels (the number of occupied radio channels) in each cell in order to observe the traffic change over the period of two years.

The developed OPNET simulation model does not capture the wireless segment of the E-Comm network that handles communications between the base stations and the mobile radio transceivers. The E-Comm network serves several thousand users and the collected activity data do not contain the location of the users during a call. Furthermore, link errors and propagation phenomena in the wireless section of the network do not affect the establishing and discarding of the calls. Hence, they do not affect the network utilization. The limited cell capacities (number of available radio channels) are the network bottlenecks. Therefore, it is important to simulate the number of occupied channels in each cell.

Mobility of radio devices and call handover are two major concerns for micro-cell cellular networks. However, they do not affect the operation of the E-Comm network. The E-Comm network is a wide-area radio network with each cell (system) covering a citywide area. Because an emergency call lasts 3.8 seconds on average, there is only a negligible probability that one radio device moves between two cells during such a short time interval. The mean value of call duration is short because the E-Comm employs the push-to-talk (PTT) mechanism for network access and because it utilizes transmission trunking rather than traditional message trunking.

The thesis includes six additional Chapters. In Chapter 2, we describe the architecture and operation of the E-Comm network. The E-Comm traffic data models are given in Chapter 3. The OPNET simulation model is presented in Chapter 4. Statistical concepts and tools employed for the traffic analysis are introduced in Chapter 5. Simulation results, statistical modelling, analysis of network performance, and a comparison of the parameters of traffic from various years are shown in Chapter 6. We conclude with Chapter 7.

CHAPTER 2: DESCRIPTION OF THE E-COMM NETWORK

We described here architecture and operation of the E-Comm network.

2.1 Architecture of the E-Comm Network

The Stanley Cup riots 1994 in downtown Vancouver, Canada revealed inadequate interoperability and communication between public emergency agencies. These emergency agencies operated on separate disconnected wireless radio networks. Hence, the government of British Columbia initiated establishing a new centralized emergency communications center for disaster coordination and public safety. It included public service to municipalities, regional, districts, the provincial and federal governments and their agencies, and emergency first responders' organizations throughout Southwest British Columbia. E-Comm is governed under the Emergency Communications Corporations Act (1997) and was incorporated September 22, 1997 under the BC Business Corporations Act [8].

The E-Comm Corporation provides wide area radio dispatching services for various emergency agencies throughout the Greater Vancouver Regional District (GVRD), the Sunshine Coast Regional District, Whistler, and Pemberton. Total of $160 million CAD has been invested in the E-Comm project. It has an annual operating budget of $45 million CAD [8].

The E-Comm public safety wireless radio network utilizes the Enhanced Digital Access Communications System (EDACS), manufactured by M/A-Com [44]. The EDACS architecture is shown in Figure 2.1. Its main elements are the central system controller (network switch), several radio repeaters sites (base stations), one or more fixed user sites (dispatch consoles), hundreds of mobile users, and a management console. EDACS is connected to the public switched telephone network (PSTN) via a Private Branch Exchange (PBX) gateway and to packet networks via the data gateway. System events and call activities are recorded by base stations and are forwarded every hour through the data gateway to the central database.

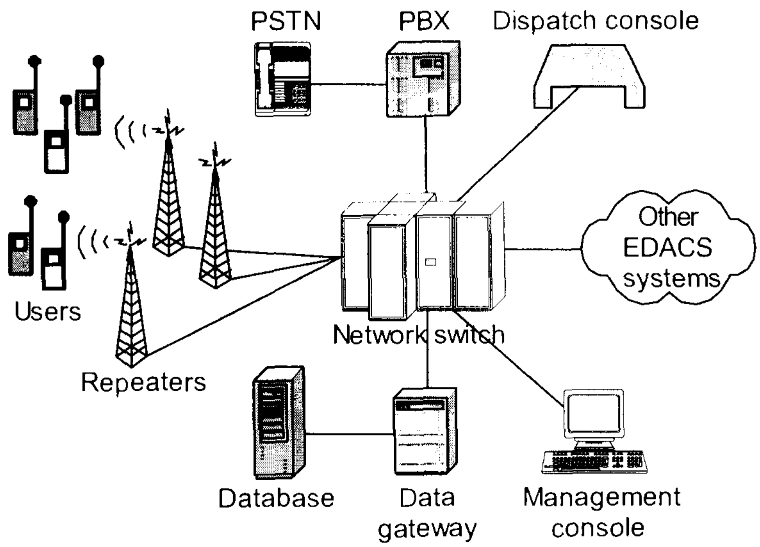

Figure 2.1: System architecture of the Enhanced Digital Access Communications System (EDACS).

The wireless section of the E-Comm network has a cellular architecture. It consists of eleven cells connected to a central switch. Each cell is covered by one or more radio repeaters (depending of cell's area) capable of transferring data using a set of frequencies. Individual cells that cover separate geographic regions in the E-Comm network are connected to the network switch by high-speed optical fibre or microwave data links. Adjacent cells are covered by distinct frequencies and there is no interference among cells.

2.2 Operation of the E-Comm Network

EDACS systems may be configured in one of the five modes: single site system, voted system, simulcast system, single channel system, and multi-site system [44].

The E-Comm utilizes simulcast configuration of EDACS systems as transmission method between the repeaters and the mobile users, which implies that all repeaters belonging to one cell use an identical set of radio carrier frequencies to transmit and receive identical audio and data information. The Simulcast systems are employed when it is necessary to use a limited number of frequencies to cover an area too large for a single repeater. For example, City of Vancouver is covered by five simulcast repeaters. The simulcast system also provides higher signal strength and by utilizing intelligent repeaters, each with distributed processing power (if one repeater fails, the other repeater controllers automatically replace the functions of the failed unit) better fault tolerance.

The number of frequencies in each cell is predefined and it determines the number of available radio channels. Each radio channel occupies one frequency. Thus, we can define the capacity of a cell as a number of its radio channels. This number also determines the maximum number of simultaneous radio transmissions (calls) in a given cell. Every radio transmission is treated as one call. Each call in a given cell occupies one radio channel. In each cell, one frequency is dedicated to the exchange of control information before, during, and after the call. Hence, the capacity of a cell (number of user channels) is one less than the number of available frequencies. Cell's capacity is determined based on the expected traffic volume. No protocol and traffic data for the control channel were available in the collected data. Hence, we only analyzed the network traffic and the utilization of user channels.

The management of frequencies (radio channels) in the E-Comm network is based on trunking. Trunking implies sharing all frequencies in a cell among all agencies instead of dedicating subsets of frequencies to individual agencies. This approach results in better utilization of radio resources and minimizes the number of radio channels necessary for matching certain Grade of Service (GoS) requirements. The E-Comm network utilizes transmission trunking (each radio transmission is treated as a separate call) rather than message trunking. Transmission trunking is 20–25% more efficient than message trunking [44]. However, there is overhead because of the high channel utilization in the transmission trunking mode (channel assigning time and channel dropping time are added to each transmission because the processes of channel assigning and channel dropping are repeated over and over for every press of the PTT button). The channel assigning and channel dropping times are 0.25 seconds and 0.16 seconds, respectively. The partial solution to this overhead is high-speed control channel in the E-Comm network, which is available all the time. The control channel supports 9.6 kbps digital signaling. This very fast channel access provides conditions for transmission trunking. With transmission trunking, a call is established (a channel is dedicated to the transmission) when a user presses the push-to-talk (PTT) button on the mobile radio transceiver. The call lasts as long as the user holds the PTT button. The call ends and the radio channel is released when the PTT button is released. Simplified schematic diagram of the general structure of a trunked radio system is shown in Figure 2.2.

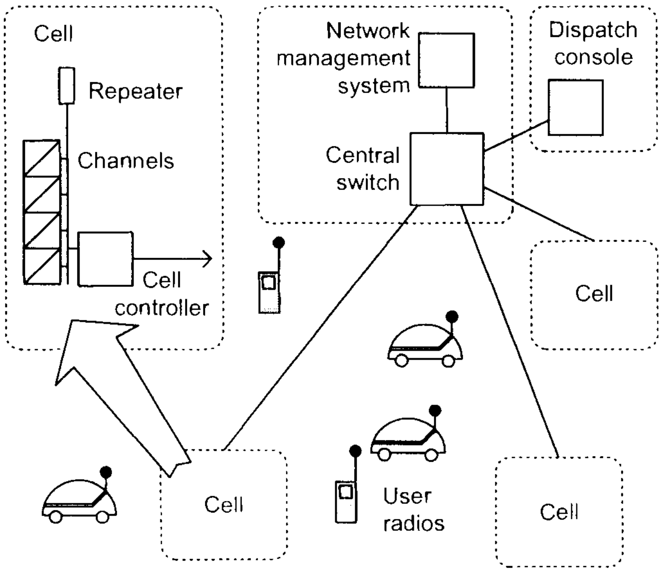

Figure 2.2: Schematic diagram of a trunked radio system.

In the E-Comm network, users are organized into talk groups that belong to agencies such as police, fire department, and ambulance. They are defined at various levels (agency level, fleet level, and sub-fleet level) for better coordination of operations. A user (radio device) may be a member of more than one talk group and may switch between talk groups dynamically. The system serves approximately 600 talk groups, consisting of a variable number of users (units) that often ranges between 20 and 150. For the sample week in 2002, members of a single talk group appeared in 2.26 cells on average. For the sample week in 2003, the average number was 2.54 cells. The most common type of a call in the network is the group call. Call recipients are members of a talk group. The advantage of this type of call is that it eliminates the need for radio users to know the target device number in order to make connection with particular user.

Usually, users call a target call group and the current members of the group without knowing device numbers.This implies a one-to-many type of conversation: a user talks to all other members in his/her talk group. Depending on the locations of the members of the talk group, a call may require one or more channels. If all members of the talk group are within one cell, the call is established using one free channel. If members of the talk group reside in several cells, network controller will allocate to the call a free channel in each destination cell. Therefore, a single call might use simultaneously several channels.

A call is established by using a push-to-talk mechanism. A user (member of a talk group) talks to other members of the talk group by pressing the push-to-talk button on the mobile radio transceiver. The central system controller then determines the locations of the talk group members and checks for availability of radio channels in every cell where the members are located. If there is at least one free channel in every cell, the caller receives an audible signal to establish the call. The one-way communication (call) lasts as long as the initiator holds the push-to-talk button. If there are no available channels in at least one destination cell with members of the talk group, the call is blocked and queued. The call is discarded (dropped) if it cannot be established after a certain period of time. In the analyzed dataset, the number of queued calls is negligible (< 0.5%) compared to the number of established calls.

Each cell has a distinct pool of frequencies. The number of frequencies determines its capacity. Cells, with their predefined and limited capacities, are main network bottlenecks. Queued and dropped calls occur due to the insufficient number of radio channels in the cells. Therefore, analyzing and modelling call traffic from each cell individually is important to determine current and predict future network performance.

Each cell covers a relatively large area (entire municipality) and the calls are relatively short (average 3.8 s [4]). This implies rare occurrence of call handover.

Table 2.1 shows the system ID, system coverage, and number of deployed channels in each system in the E-Comm network as of December 2003.

Table 2.1: Number of channels deployed in each of E-Comm cells (December 2003).

System ID	Coverage	Number of channels
1	Vancouver	13
2	Burnaby	9
3	Maple Ridge	7
4	Langley	6
5	Seymour	6
6	Port Coquitlam	8
7	Richmond	7
8	Mission	5
9	Surrey	9
10	South Surrey	8
11	Bowen Island	4

CHAPTER 3: TRAFFIC DATA MODELS

In this Chapter, we introduce traffic data models and the procedure for pre-processing the original (raw) E-Comm traffic data. Activity data from the deployed network recorded by E-Comm consist of records of network events: established, queued, and discarded calls, as well as talk group dynamics. All network events occurred in the E-Comm network are recorded by base stations and forwarded every hour through the data gateway to the central database that contains call activity information from the whole system. To analyze the behavior of the E-Comm network, traffic data of three consecutive years were compared.

3.1 Traffic Data Pre-processing

Three sets of call traffic data from the E-Comm network were available for our analysis. The analyzed data sets span 2 days in 2001 year (2001-11-01 to 2001-11-02), 30 days in 2002 year (2002-02-09 to 2002-03-10), and 92 days in 2003 year (2003-03-01 to 2003-05-31) of call traffic data. As example of how big the volume of the original data set is: the size of the database of 2003 call traffic data is 6 Gbytes, with 44,786,489 records for the 92 days of data. It consists of 92 event log tables, each containing records of one day's events. The large size of the data set was one of the difficulties in our data analysis. The original data were in MS Access format, and we converted the data to plain text files and imported the records into a MSSQL database server on a Linux platform for further processing.

Table 3.1 represents the time span and the number of calls in 2001 and for one-week data in 2002 and 2003 in each dataset.

Table 3.1: Time span and number of calls in the traffic traces.

Trace (dataset)	Time span	No. of calls
2001	November 1–2, 2001	110,348
2002	March 1–7, 2002	370,510
2003	March 24–30, 2003	387,340

The E-Comm database contains event log tables recording all events occurred in the network. Each row in the original data set represents one event that occurred in the E-Comm network. A sample of the E-Comm network event log table (raw traffic data) is given in Table 3.2. It consists of twenty six types of fields.

Descriptions of twenty six types of the data fields appearing in the E-Comm event log table (traffic data) are:

Event_UTC_At: Call arrival timestamp of the event with given granularity of 3ms

Duration_ms: Call holding time (call duration in ms with given granularity of 10ms)

System_Id: The identification of the system (cell) in which a call occurred (range from 1 to 11)

Channel_Id: The identification of the channel in which a call was established

Table 3.2: A sample of the E-Comm network event log (raw traffic data).

Event_UTC_At		Duration_ms	System_Id	Channel_Id
2003-03-10 00:10:27.280		2480	1	C
2003-03-10 00:10:27.300		2470	2	C
2003-03-10 00:10:27.790		2490	6	C
2003-03-10 00:10:27.877		2490	8	C
Caller	Callee	Queue_Depth	Network_Id	Node_Id
A	B	NULL	1	33
A	B	NULL	1	33
A	B	NULL	1	33
A	B	NULL	1	33
Call_Type	Call_State	Slot_Id	Call_Direction	Voice_Call
0	0	NULL	0	1
0	0	NULL	0	1
0	1	NULL	100	1
0	1	NULL	100	1
Digital_Call	Interconnect_Call	Multi_System_Call	Confirmed_Call	Msg_Trunked_Call
1	0	0	0	0
1	0	0	0	0
1	0	0	1	0
1	0	0	1	0
Preempt_Call	Primary_Call	Queue_Pri	MCP	Caller_Bill
0	1	NULL	1	0
0	1	NULL	1	0
0	1	NULL	0	0
0	1	NULL	0	0
Callee_Bill	Reason_Code			
0	0			
0	0			
0	0			
0	0			

Caller:	The identification of a radio device that initiates a call. It is the caller's ID, ranging from 1 to 16,000. The first 2,000 caller's IDs are dedicated to either talk groups or individual users and rest of them are assigned to talk groups only.
Callee:	The identification of a radio device that receives a call in same range as caller's ID
Queue_Depth:	The number of calls waiting in the queue at the event time instance
Network_Id:	The network identification (constant equal to "1")
Node_Id:	The identification of the network node (constant equal to "33")
Call_Type:	The type of the call (group/individual/emergency/group set/system all/Morse code/test/paging/scramble data/sys login/start emergency/cancel emergency)
Call_State:	The state of the call (channel assign/channel drop/key/ un-key/digits/over digits/ queue/busy/deny/convert to callee)
Slot_Id:	The constant equal to "NULL"
Call_Direction:	Indicator of the call direction (making or receiving a call)
Voice_Call:	The flag indicating if the call contains voice information

Digital_Call:	The flag indicating if the call is from digital device or analog device
Interconnect_Call:	The flag indicating if it is call connecting Enhanced Digital Access Communications System (EDACS) to Public Switched Telephone Network (PSTN)
Multi_System_Call:	The flag indicating if it is a multi-system call and it is only set in the event of call drop
Confirmed_Call:	The flag indicating if it is a confirmed call (each member of a talk group has to confirm the call before beginning of the conversation)
Msg_Trunked_Call:	The flag indicating if the call is message trunking or transmission trunking
Preempt_Call:	The flag indicating if it is a pre-empt call (it has higher queue priority)
Primary_Call:	The flag indicating if it is a real call or group set signal
Queue_Pri:	The priority number of a call in a queue
MCP:	Indicator of multiple channel partition
Caller_Bill	Indicator, who (caller or callee) will pay for the call. It is set to 1 if caller will pay for the call
Callee_Bill:	Indicator, who (callee or caller) will pay for the call. It is set to 1 if callee will pay for the call

Reason_Code: The error reason code number gives additional information about errors if any appear during the call.

One call could usually generate two or more events and hence there are a number of redundant records in the raw data set. For example, calls involving multiple systems (Multi_System_Call) could generate a channel-assigning event and a channel-dropping event in all systems. In that case, in addition to records with call state = 0 (the call assignments events) in the data base may exist and same number of records with call state = 1 (the call drop events) that refer to the very same call. Since this call has already records in the database (the records with call state = 0), the corresponding records with call state = 1 are redundant. Notice that each call drop event already has a corresponding call assignment event in the database while the reverse does not hold. In addition, there are other redundant records, such as records having call type = 100 or records with duration = 0. The complete call type table is given in Appendix A.1. Since duplicate records of same call are not of interest for our analysis, they are deleted. We also have removed the records with channel id = 0 because it is the control channel and the traffic data from this channel were not of particular interest for our research. Further, certain fields in the database have NULL value (Queue_Depth, Queue_ Pri, and Slot_Id), while others have identical values (Network_Id, Node_Id, Caller_Bill, and Callee_Bill), so the columns that correspond to these records are erased from data sets.

The nine fields that capture the user's behaviour and network traffic are of particular interest to our study: Event_UTC_ At, Duration_ms, System_Id, Channel_Id, Caller, Callee, Call_Type, Call_State, and Multi_System_Call, while other fields (Call_Direction, Interconnect_Call, Digital_Call, Voice_Call, MCP, Confirmed_Call, Msg_Trunked_Call, Preempt_Call, Primary_Call and Reason_Code) are not useful to our analysis and are disregarded.

A sample of the cleaned traffic data (2003-03-01), after reducing the database dimension to nine is shown in Table 3.3.

Table 3.3: A sample of the processed call traffic data.

No	Time (hh:mm:ss)(ms)	Call Duration (ms)	System Id	Channel Id	Caller	Callee	Call Type	Call State	Multi System Call
1	00:00:00 30	1340	1	12	13905	401	0	0	0
6	00:00:00 489	1350	7	4	13905	401	0	0	0
29	00:00:03 620	7550	2	7	13233	249	0	0	0
31	00:00:03 760	7560	1	3	13233	249	0	0	0
37	00:00:04 260	7560	7	6	13233	249	0	0	0
38	00:00:04 340	7560	6	6	13233	249	0	0	0

As shown in Table 3.3, a multi-system call is recorded by multiple entries. By observation of the caller, callee, and call duration information, we may conclude that records 1 and 6 represent one group call from caller 13905 to callee 401, involving systems 1 and 7 and lasting 1,350 ms. Records 29, 31, 37, and 38 represent a group call from caller 13233 to callee 249, involving systems 2, 1, 7, and 6.

The data are aggregated from the distributed database of the individual network management systems. The transmission latency and glitches in the distributed database system cause that multiple records with identical caller id and callee id and similar call duration fields might represent one single group call in the database. For example, records 1 and 6 in Table 3.3 have 10 ms difference in call duration field although they represent one single group call. It is desirable to combine these records into one row which represent one call. We used 10ms difference in call duration as an empirical choice when combining the multiple records.

After the data preprocessing, size of the data base was reduced for more than half and the number of records in the database was reduced to only 19% of the original records after removing irrelevant data fields for our research.

3.2 OPNET Data Model

A data model for OPNET simulations was created based on two weeks of activity data from the E-Comm records. We compare the network performance during similar time periods in the two years. The 2002 sample data span the week between 0:00 on February 25, 2002 and 24:00 on March 3, 2002. The 2003 sample data span the week from 0:00 on March 10, 2003 to 24:00 on March 16, 2003. Timestamps end in either 0 or 7 (e.g., 2003-03-10 0:10:27.280 and 2003-03-10 0:10:27.877) due to the limitation of the "datetime" data type in the MS SQL Server [51] that is used by E-Comm to record events. Hence, the resolution of the timestamps is at least 10 ms. The resolution of the call durations is 10 ms. In order to analyze the utilization of radio channels in individual cells, the sample data were aggregated. The data model, formatted as a trace file, was used as input for OPNET trace-driven simulations. Figure 3.1 shows preparation of the traffic trace file from the E-Comm activity data. From the sample data, we extracted only the records relevant to established voice calls, indicating the caller, the called talkgroup (callee), the time of a call (timestamp), and how long a channel in a given cell was occupied (duration). An excerpt from the 2003 sample data showing only the relevant fields is given in Table 3.4. Each user has a unique user ID and each talk group has its unique identification number. In this example, the four rows correspond to one call. The call began at approximately 0:10:27 on March 10, 2003. It lasted 2,490 ms and involved cells 1, 2, 6, and 8. (To maintain confidentiality of the data, the ID's of the caller, callee, and caller agency were labelled A, B, and C, respectively.) The caller agency ID identifies the corresponding public safety agency. This field is not used in the OPNET simulations. It is used by WarnSim [2], [43] to analyze traffic emanating from individual agencies.

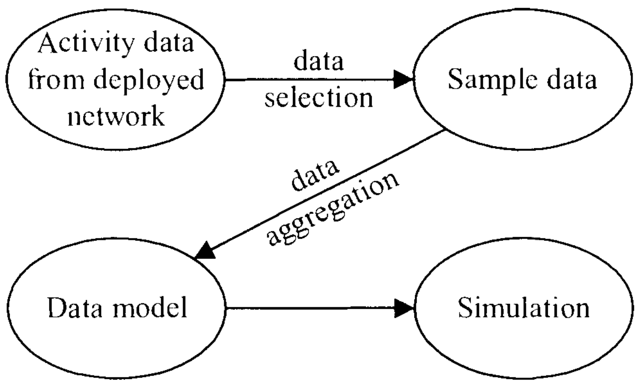

Figure 3.1: Preparing the traffic trace file from the E-Comm activity data.

Table 3.4: Excerpt from the 2003 sample data.

Timestamp	Duration (ms)	Caller	Callee	Cell	Caller agency
2003-03-10, 0:10:27.280	2,480	A	B	1	C
2003-03-10, 0:10:27.300	2,470	A	B	2	C
2003-03-10, 0:10:27.790	2,490	A	B	6	C
2003-03-10, 0:10:27.877	2,490	A	B	8	C

WarnSim is publicly available wide area radio network simulator. It works under Windows platforms with .NET framework 1.0 (and up) support.

A call can be represented by one or more rows in the sample data. The number of rows represents the number of cells where the call terminates. A call in the deployed network is uniquely identified by four fields: timestamp, duration, caller (ID of the user who initiated the call), and callee (ID of a talkgroup that receives the call). Nevertheless, timestamps and durations corresponding to a single call differ due to discrepancies in the records (sample data), as shown in Table 3.4. For the data model, we arbitrarily chose the smallest timestamp. The largest call duration was chosen in order to simulate the worst-case scenario. We also modified the format of the timestamp. The original timestamp represented the date and time of the beginning of a call. For simulation purposes, it was convenient to express the timestamp as a difference between the original timestamp and an arbitrary reference time. The reference times were chosen to be 0:00 on February 25, 2002 and 0:00 on March 10, 2003 for the 2002 and 2003 data models, respectively. In order to create trace files used in simulations, we modified the sample data so that one row corresponds to one call. As a result, one record in the OPNET model trace file (the data model) that corresponds to the four rows of data shown in Table 1 is: {627280, 2490, 1, 2, 6, 8}, where 627,280 is the timestamp (ms) calculated from reference time instant, 2,490 is the duration of the call (milliseconds), and the remaining numbers are the cell IDs where the call terminated.

3.3 Traffic Data Model used for Statistical Modelling

Traffic data from E-Comm consist of records of network events, such as established, queued, and dropped calls. Each established call is identified by its timestamp, duration, caller, callee, and destination cell(s). By neglecting the mobility and call handover components, the analyzed data may be captured by two random processes: call arrival process and call duration process.

From the traffic data, we create traces of call holding times (call durations) and call inter-arrival times (differences between two successive timestamps). We analyze traces from the cell covering Vancouver because it is the busiest cell and handles the majority of the calls. It also has the largest number of available radio channels and a sufficient capacity so that congestion and call queuing rarely occur.

Analyzed traffic traces (datasets) span various periods during three years: 2001, 2002, and 2003. We determine the number of calls in every one-hour interval of each dataset in order to identify the busiest hours. Our analysis focuses on call holding and call inter-arrival times from the five busiest hours in each dataset. Analysis of busy hour traffic is typical for circuit-switched networks because they are designed to satisfy certain Grade of Service (GoS) requirements regarding the frequency of occurrence and duration of call queuing during periods of high utilization [41]. The number of calls during the five busiest hours in each dataset is shown in Table 3.5.

Table 3.5: **Five busiest hours in the traffic traces from 2001, 2002, and 2003 with the corresponding number of calls.**

2001		2002		2003	
Day/hour	No.	Day/hour	No.	Day/hour	No.
02.11.2001 15:00–16:00	3,718	01.03.2002 04:00–05:00	4,436	26.03.2003 22:00–23:00	4,919
01.11.2001 00:00–01:00	3,707	01.03.2002 22:00–23:00	4,314	25.03.2003 23:00–24:00	4,249
02.11.2001 16:00–17:00	3,492	01.03.2002 23:00–24:00	4,179	26.03.2003 23:00–24:00	4,222
01.11.2001 19:00–20:00	3,312	01.03.2002 00:00–01:00	3,971	29.03.2003 02:00–03:00	4,150
02.11.2001 20:00–21:00	3,227	02.03.2002 00:00–01:00	3,939	29.03.2003 01:00–02:00	4,097

Figure 3.2 shows a time series of the call traffic between 22:18 and 22:19 on March 26, 2003 (one minute of the busiest hour in the 2003 dataset). The horizontal axis shows the timestamps of the call. The vertical axis shows the call holding times. Call inter-arrival times are observed as time intervals between two successive calls.

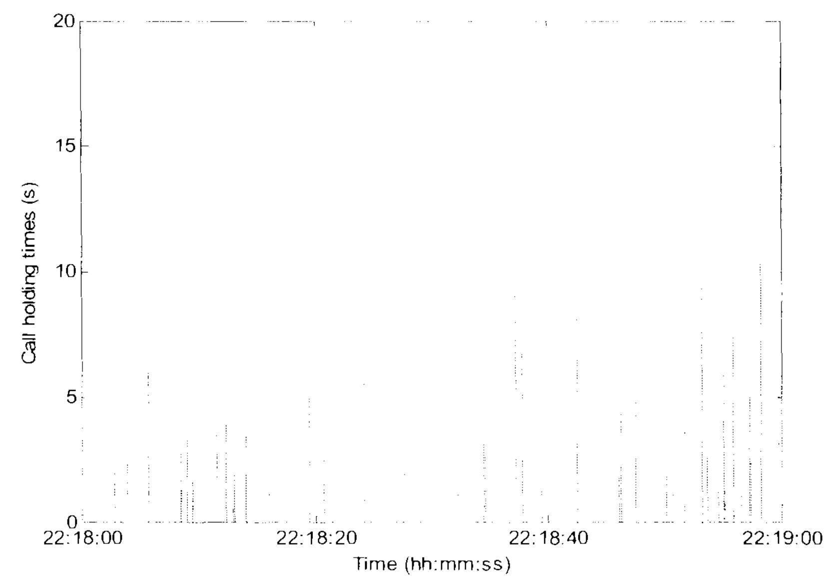

Figure 3.2: Time series of the one-minute interval between 22:18 and 22:19 from the 2003 busiest hour traffic.

CHAPTER 4: THE OPNET SIMULATION MODEL

We used the OPNET [42] network simulator to analyze the E-Comm network performance. In this study, we consider only the circuit-switched network segment that carries user traffic between mobile users. OPNET network models have a hierarchical architecture with three layers: network, node, and process. The network topology represents the top layer of the OPNET network model and consists of network nodes. Nodes consist of interconnected modules that perform defined tasks and exchange information using packet streams and statistical wires. The functionality of each module is defined using a process model. The process model is created by using a finite state machine (FSM) approach and the OPNET specific functions. The role of process model is to mimic behaviour of a system in response to events. This is accomplished through states and transitions that graphically specify activities of a process in response to the particular events. Each state of a process model is defined by using C/C++ code.

4.1 OPNET Network Model

The OPNET network model, shown in Figure 4.1, consists of a central switch and eleven cells located in various regions of the Greater Vancouver Regional District. The cells are connected to the central switch via point-to-point simplex links. In the deployed system, after call establishment, voice information flows from the originating cell to the central site and then to the destination cell(s). The call establishment procedure is performed by exchanging information over the control channel. The data model does not contain information about the originating cell of a call, and, therefore, the traffic in the OPNET model is generated in the central site and is then sent to the corresponding cells.

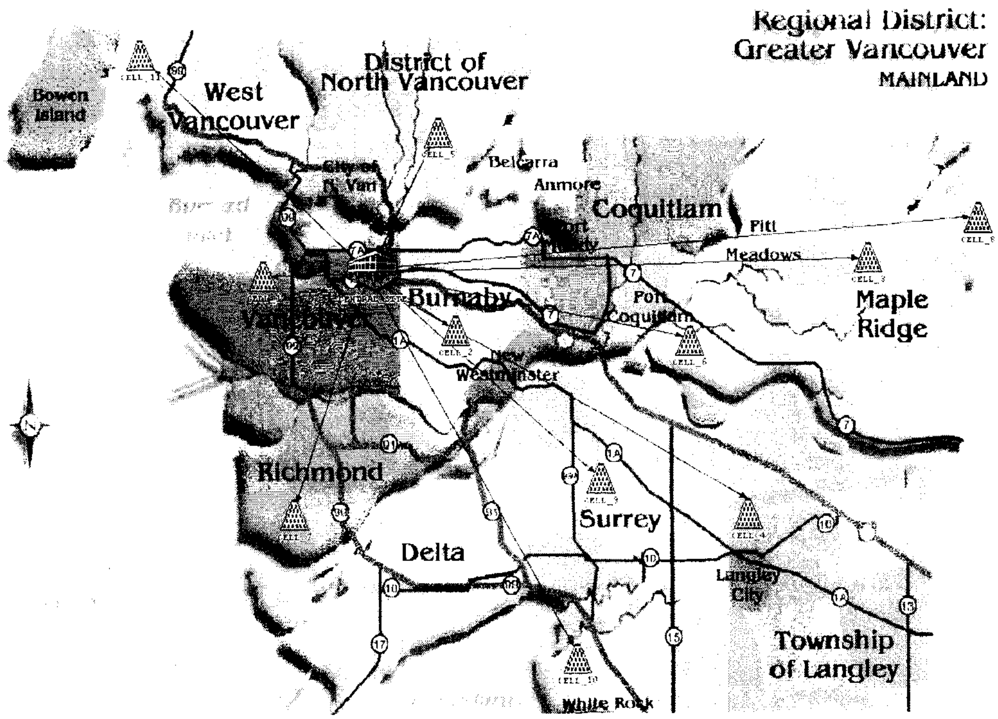

Figure 4.1: OPNET model of the E-Comm network. The network model consists of a central switch located in East Vancouver and eleven cells covering various municipalities of the Greater Vancouver Regional District. The cells are connected to the central site via point-to-point links.

Each link has a number of channels equal to the number of frequencies available in a cell. The cell capacities (number of user channels) are shown in Table 4.1.

Table 4.1: Number of user channels per cell. The only difference between the 2002 and 2003 data sets is the capacity of cell 9.

Cell	1	2	3	4	5	6	7	8	9	10	11
Channels (2002)	12	7	4	5	3	7	6	4	6	6	3
Channels (2003)	12	7	4	5	3	7	6	4	7	6	3

One occupied frequency in a cell corresponds to one busy channel in the link that connects the cell with the central site. Therefore, during simulations, we recorded the number of used frequencies (the instantaneous utilization of radio resources) by monitoring the utilization of each point-to-point link (number of occupied channels). All channels in a link have an identical bit rate (arbitrarily chosen to be equal to 1,000 bits per second). In the OPNET model, calls are represented by packets. When a call is forwarded to a cell, the central site generates a packet of 1,000 × *CD* bits, where *CD* is the duration of the call (in seconds). The packet is sent to an idle link channel that connects the central site and the cell. The channel in the corresponding link will be occupied *CD* seconds, starting from the instance when the call is established.

4.2 OPNET Node and Process Models

We created OPNET node and process models for the elements of the E-Comm network. We used standard OPNET process models for point-to-point transmitters, point-to-point receivers, and packet sinks.

The OPNET network model consists of the central site node and the eleven cell nodes. The node model of the central site (network switch) is shown in Figure 4.2. Its functions are reading the trace file, generating packets that correspond to calls, sending the packets to appropriate cells, and collecting statistics. These functions are implemented in the modules that constitute the central site node model: *source*, *dispatcher*, *channel_selector*, and *tx* (point-to-point transmitter). There is one *source* module, one *dispatcher* module, and eleven pairs of *channel_selector* and *tx* modules (one pair for each cell). The *source* module reads the trace file and forwards to the *dispatcher* module the information about the calls to be established (call duration and destination cells). The central module in the OPNET node model is the *dispatcher*. Its process model is shown in Figure 4.3. It consists of four states. Initialization of the statistics that are collected during the simulation is performed in the *init* state. After the *init* state, the process proceeds to the *idle* state. When the *dispatcher* receives a notification from the *source* that a call is to be established, it proceeds to the *call* state. In this state, it checks for availability of free channels in the cells and it decides whether or not the call could be established. If the call can be established, the process creates a packet of a length proportional to the duration of the call and forwards it to the corresponding *channel_selector* module(s). If the call cannot be established, the number of discarded calls is updated and the packet that corresponds to the call is destroyed. (The model does not support call queuing: blocked calls are discarded immediately.) After the *call* state, the process returns to the *idle* state. The *dispatcher* is connected to each transmitter *tx* by statistical wires that monitor the channel occupancy in each link. One statistical wire monitors a single channel. When the *dispatcher* receives a notification that the status of a channel has changed, it proceeds to the *calc_stat* state. There, the values of the collected statistics are updated and the process returns to the *idle* state. Each *channel_selector* module registers free and occupied channels in its connected link. When a packet from the *dispatcher* arrives, the *channel_selector* sends the packet via one of the free channels and marks the channel as busy. When a cell receives a packet, which is equivalent to a call being completed, the *channel_selector* marks the corresponding channel as free.

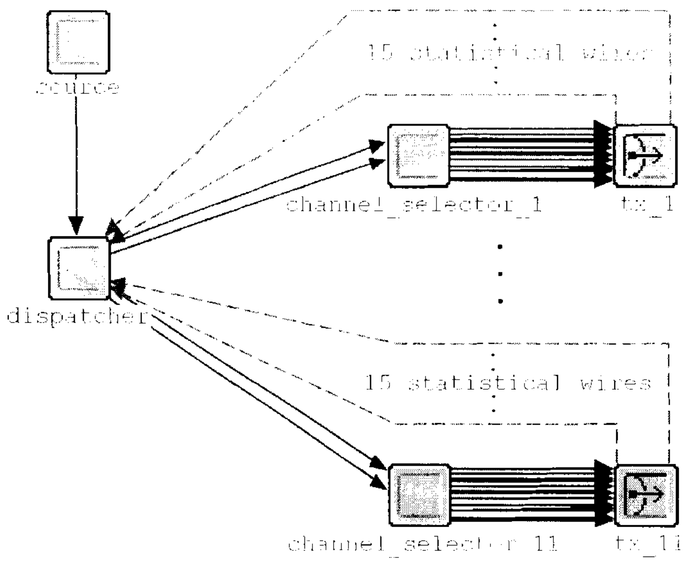

Figure 4.2: OPNET node model of the central switch.

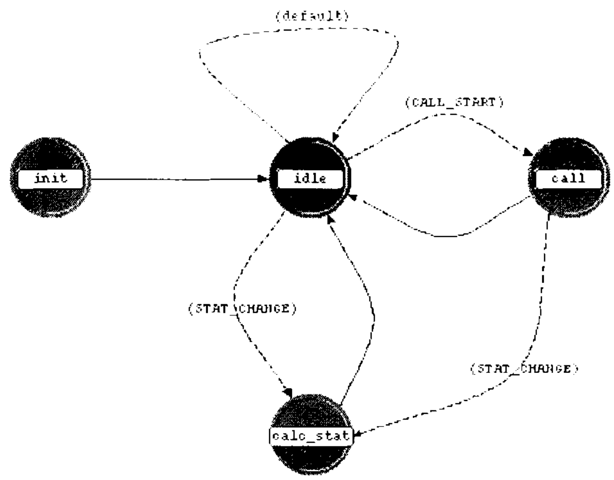

Figure 4.3: OPNET process model of the dispatcher module in the central site node model.

The node model of a cell is shown in Figure 4.4. It consists of a point-to-point receiver *rx*, a *receiver* module, and a *sink*. When a packet arrives, the *receiver* module notifies the corresponding *channel_selector* in the central site of the free channel in the link and sends the packet to the *sink*.

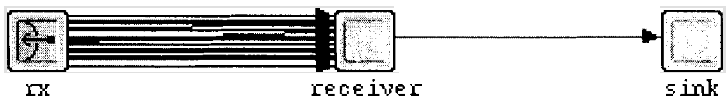

Figure 4.4: OPNET node model of a cell.

CHAPTER 5: STATISTICAL CONCEPTS AND ANALYSIS TOOLS

Statistical processes possess two important characteristics: probability distribution and autocorrelation. The probability distribution characterizes the probability that the outcomes of the process (random variables) are within a given range of values. It is expressed through probability density and cumulative distribution functions. Probability density functions show the probability of occurrence of a certain value or range of values. Cumulative distribution functions express the probability that the variable will not exceed specific values. The autocorrelation function measures the dependence between two outcomes of the process. In general, it is a function of the time instances of the two outcomes. If the process is wide-sense stationary, its autocorrelation function depends only on the difference (lag) between the time instances of the outcomes.

The traffic in the E-Comm network is characterized by two processes: call arrival and call holding processes. Outcomes of the call arrival and call holding processes are the sequences of call inter-arrival and call holding times, respectively. Investigating the statistical properties (probability distributions and the autocorrelations) of these processes is important for deriving an appropriate traffic model and employing the model for determining network performance. Choosing the statistical distribution that best fits the data is performed by comparing the distribution of the data with several known distributions and employing the Kolmogorov-Smirnov (K-S) goodness-of-fit test. The autocorrelation of the data is examined by plotting the autocorrelation functions and testing whether the data exhibit long-range dependence.

5.1 Long-Range Dependence (LRD)

For mathematical simplicity, it is often assumed that a process is wide-sense stationary and uncorrelated, or that its autocorrelation function is zero for non-zero lags. This assumption does not hold for all processes. Often, there is a certain correlation structure that cannot be neglected. A class of processes with non-negligible autocorrelations is the family of long-range dependent (second-order self-similar) processes [45], [46]. Long-range dependence is defined as a non-summability of the autocorrelation function r(k) of a wide-sense stationary process X(n), n = 1, 2, 3, ... The autocorrelation function r(k) of an LRD process is modeled as a hyperbolically decaying function

$$r(k) = c_r k^{-(2-2H)}, \; k \to \infty, \tag{1}$$

where c_r is a positive constant and H ($0.5 \leq H < 1$) is the Hurst parameter. The power spectral density (PSD) f(v) of X(n) satisfies

$$f(v) = c_f |v|^{-\alpha}, \; |v| \to 0, \tag{2}$$

where c_f is a positive constant and α is the scaling exponent [47]. For LRD processes

$0 < \alpha < 1$ and the relationship between H and α is linear:

$$H = 0.5(1+\alpha). \tag{3}$$

The Hurst parameter H measures the degree of LRD of a process. Values of $H \approx 1$ imply strong LRD (strong correlations between outcomes of the process that are far apart). For uncorrelated processes H = 0.5.

5.2 Wavelets and Wavelet-Based Estimator of H

The discrete wavelet transform (DWT) of a signal X(t) is given by the inner product:

$$d(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt \,, \tag{4}$$

where d(j, k) is the wavelet coefficient at octave j and time k and

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-j}t - k\right),\ j \in Z^{+},\ k \in Z \,, \tag{5}$$

is the basis function called a wavelet. It is obtained by scaling (by a factor of 2^{-j}) and translating (by k time units) of an adequately chosen mother wavelet ψ [47]. The mother wavelet possesses two important properties. It is an oscillating function (its mean value is zero). Furthermore, most of its energy is concentrated within limited time interval and limited frequency band. The signal X(t) is represented as a weighted sum of wavelets:

$$X(t) = \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k)\psi_{j,k}(t) \,. \tag{6}$$

The discrete wavelet transform (DWT) captures signals over various time scales. Wavelets' scale invariance makes the discrete wavelet transform (DWT) suitable for analyzing properties that are present across a range of time scales, such as LRD. Furthermore, the existence of low computational cost algorithms for implementing the discrete wavelet transform (DWT) makes DWT a popular tool for signal analysis.

The wavelet-based Hurst parameter estimator is based on the shape of the power spectral density (PSD) function (2) of the LRD signal X(t). It has been shown [47] that the power-law behaviour of the PSD implies the following relationship between the variance of the wavelet coefficients and the octave j:

$$E\left\{d(j,k)^2\right\} = 2^{j\alpha} c_f C(\alpha,\psi), \tag{7}$$

where the average is calculated for various k, α is the scaling exponent, and C(α,ψ) depends on the mother wavelet, but does not depend on j. When the mother wavelet is suitably chosen [47], $E\left\{d(j,k)^2\right\}$ is a sample mean of $d(j,k)^2$ calculated over all k's:

$$E\left\{d(j,k)^2\right\} = \frac{1}{n_j}\sum_{k=1}^{n_j} d(j,k)^2, \tag{8}$$

where n_j is the number of coefficients available at octave j. The plot of $\log_2 E\left\{d(j,k)^2\right\}$ vs. j is called a logscale diagram. Linear relation with a slope α (0 < α < 1) between $\log_2 E\left\{d(j,k)^2\right\}$ and j for a range of octaves, including the coarsest, indicates presence of LRD. Therefore, α is obtained by performing linear regression of $\log_2 E\left\{d(j,k)^2\right\}$ on j over a range of octaves. The Hurst parameter is obtained from relation (3).

We employed the publicly available MATLAB code [48] to estimate H. In the analysis, we used the wavelet Daubechies3. The Daubechies wavelets are a family of orthogonal wavelets. They utilize the existence of two orthogonal bases (related by an orthogonal transformation) for a model space. The first basis function is used to sample the data by averaging on small scales. It is the scaling function. The orthogonal transformation is generated by filtering the coefficients of the first basis function into equal number of high and low frequency samples. The high frequency samples are stored as coefficients for later reconstruction of the signal and they represent the second basis function called the mother wavelet. The signal is decomposed in this manner further and further and the remaining low frequency samples are further divided into high and low frequency parts. This process is repeated until there is only one low frequency coefficient remaining. Wavelets are characterized by a number of vanishing (zero) moments equal to half the number of coefficients. The index number (Daubechies3) corresponds to the number of coefficients. Wavelets ability to approximate polynomial features of a signal is represented by the number of zero moments [47].

5.3 Test for Time Constancy of the Scaling Exponent α

LRD processes are, by definition, wide-sense stationary. However, they possess certain characteristics that make them seem non-stationary. LRD processes exhibit high variability [49] and there are relatively long on and off periods. An important issue is how to distinguish between wide-sense stationary processes with LRD and inherently non-stationary processes.

An approach to determine whether a process is LRD or non-stationary is to test if the scaling exponent α is constant over the examined time series [49]. Time constancy of α is also important because the wavelet-based estimator may produce unreliable estimates when applied to time series with variable α. The test for time constancy of α [48] divides the time-series into m blocks of equal lengths and estimates α for each block. The estimates are compared and a decision is made whether or not α can be considered constant over the duration of the entire time series.

5.4 Kolmogorov-Smirnov Test

Kolmogorov-Smirnov (K-S) goodness-of-fit test is employed to determine the best fit among several distributions [50]. The null hypothesis H_0 implies that data samples follow a given distribution. The alternative hypothesis H_1 states the opposite. The purpose of the test is to check whether to accept or reject the null hypothesis H_0 and to quantify the decision. The approach of the K-S test is to examine whether the empirical distribution of a set of observations (empirical cumulative distribution function) is consistent with a random sample from an assumed theoretical distribution. The empirical cumulative distribution function E_N is defined as a step function (with step size 1/N) of N ordered data points $Y_1, Y_2, \ldots, Y_N$:

$$E_N = \frac{n(i)}{N}, \tag{9}$$

where *n(i)* is the number of data samples with values smaller than Y_i.

The decision whether or not to accept the null hypothesis H_0 is based on the value of the test statistic k, defined as the maximum difference over all data points:

$$k = \max_{1 \le i \le N} \left| F(Y_i) - \frac{i}{N} \right|, \tag{10}$$

where F is the cumulative distribution function of the assumed distribution. It means that for each data point, the K-S test compares the proportion of values less than that data point with the number of values predicted by the assumed distribution. The null hypothesis is accepted if the value of the test statistic is lower than the critical value. Three additional parameters play an important role in analyzing the test results. The significance level α (default value equals 0.05) determines that the null hypothesis is rejected α percent of the times when it is in fact true. It defines the sensitivity of the test. Smaller values of α imply larger tolerance (larger critical values). The second parameter *tail* specifies whether the K-S performs a two sided test (default) or alternative tests from one or other side. The third parameter is the observed *p-value* that reports the probability level on which the difference between distributions (test statistics) becomes significant. If $p \leq \alpha$, the test rejects the null hypothesis. Otherwise, the null hypothesis is accepted. Parameters α and *tail* are input parameters and the p-level is one of the test results. If the test returns a non-number for the critical value, then the decision to accept or reject the null hypothesis is based only on the p-value [50].

A difficulty in applying goodness-of-fit tests is that results depend on the sample size [50]. It is not uncommon for the test to reject the null hypothesis when large datasets are tested. A solution is to perform the test on randomly chosen subsets of data.

CHAPTER 6:
ANALYSIS OF DATA

We here observed the activity data from the deployed network managed by E-Comm over a period of three consecutive years, 2001, 2002, 2003, and we here analyzed the variation in time providing a comparison between them.

6.1 Network Activity

We here simulated the network activity during the two sample weeks: February 25 – March 3, 2002 and March 10 – 16, 2003 using the OPNET [42] simulation tool.

The OPNET simulation results are shown in Figure 6.1. The horizontal axes represent time (common to all graphs). In Figure 6.1, the first tick (marked 0d) corresponds to 0:00 on February 25, 2002, while the last tick (marked 7d) corresponds to 24:00 on March 3, 2002. In Figure 6.2, the first tick (marked 0d) corresponds to 0:00 on March 10, 2003, while the last tick (marked 7d) corresponds to 24:00 on March 16, 2003. The numbers of occupied radio channels during the simulated week in 2002 are shown by the first eleven graphs in Figure 6.1. They are named "Occupied channels [*i*]", where *i* corresponds to the cell ID (1 to 11). The second graph from the bottom in Figure 6.1 corresponds to "Discarded calls". Its value is equal to 1 when a call is discarded. The total number of discarded calls over time is shown on the bottom graph in Figure 6.1, labelled "Cumulative discarded calls". The same holds for graphs shown in Figure 6.2.

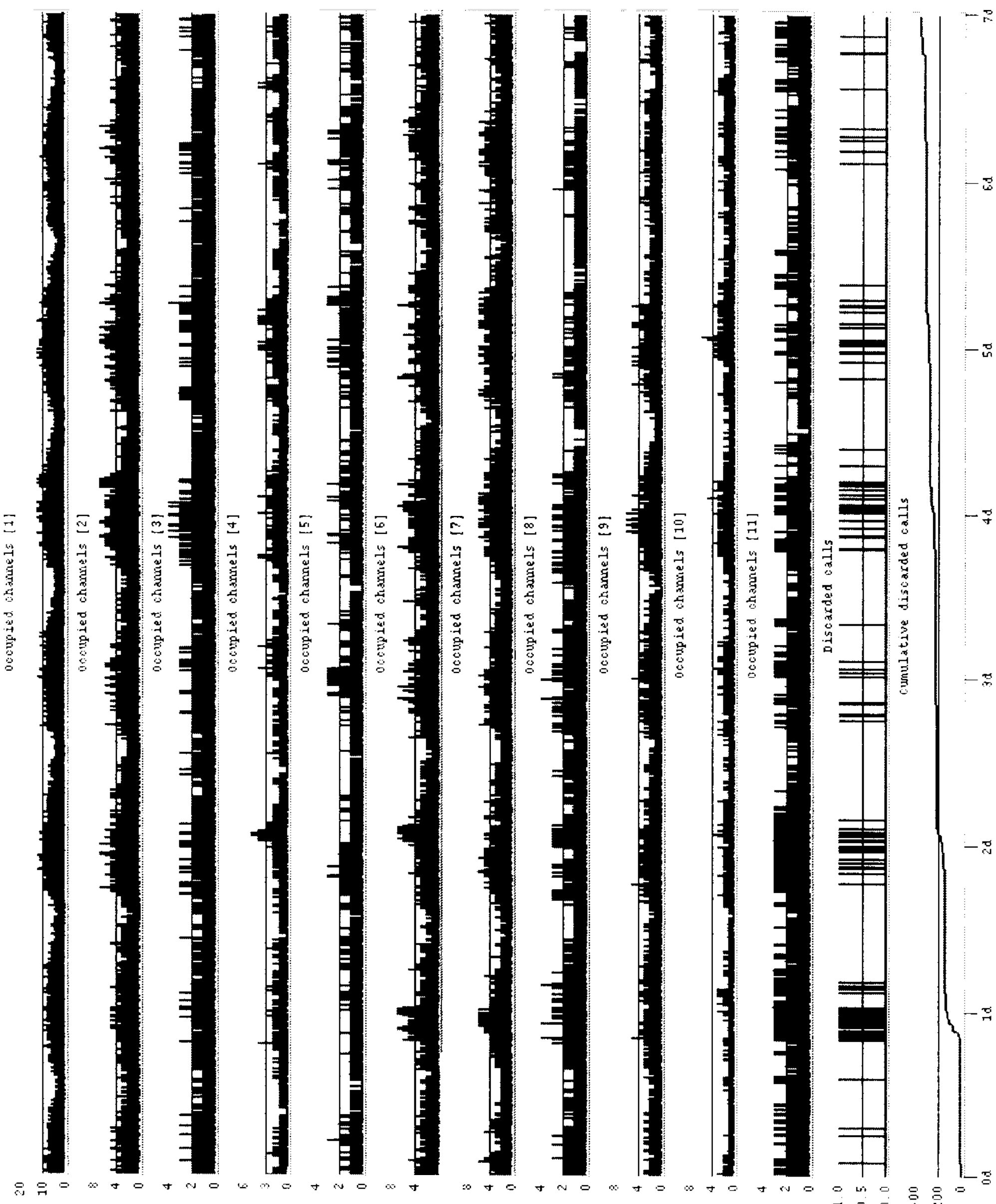

Figure 6.1: OPNET statistics collected during the simulation of network activity in 2002. "Occupied channels" graphs [1] to [11] show the utilization of each cell (number of occupied radio channels). "Discarded calls" graph indicates the time instances when calls are discarded. "Cumulative discarded calls" graph shows the cumulative number of discarded calls.

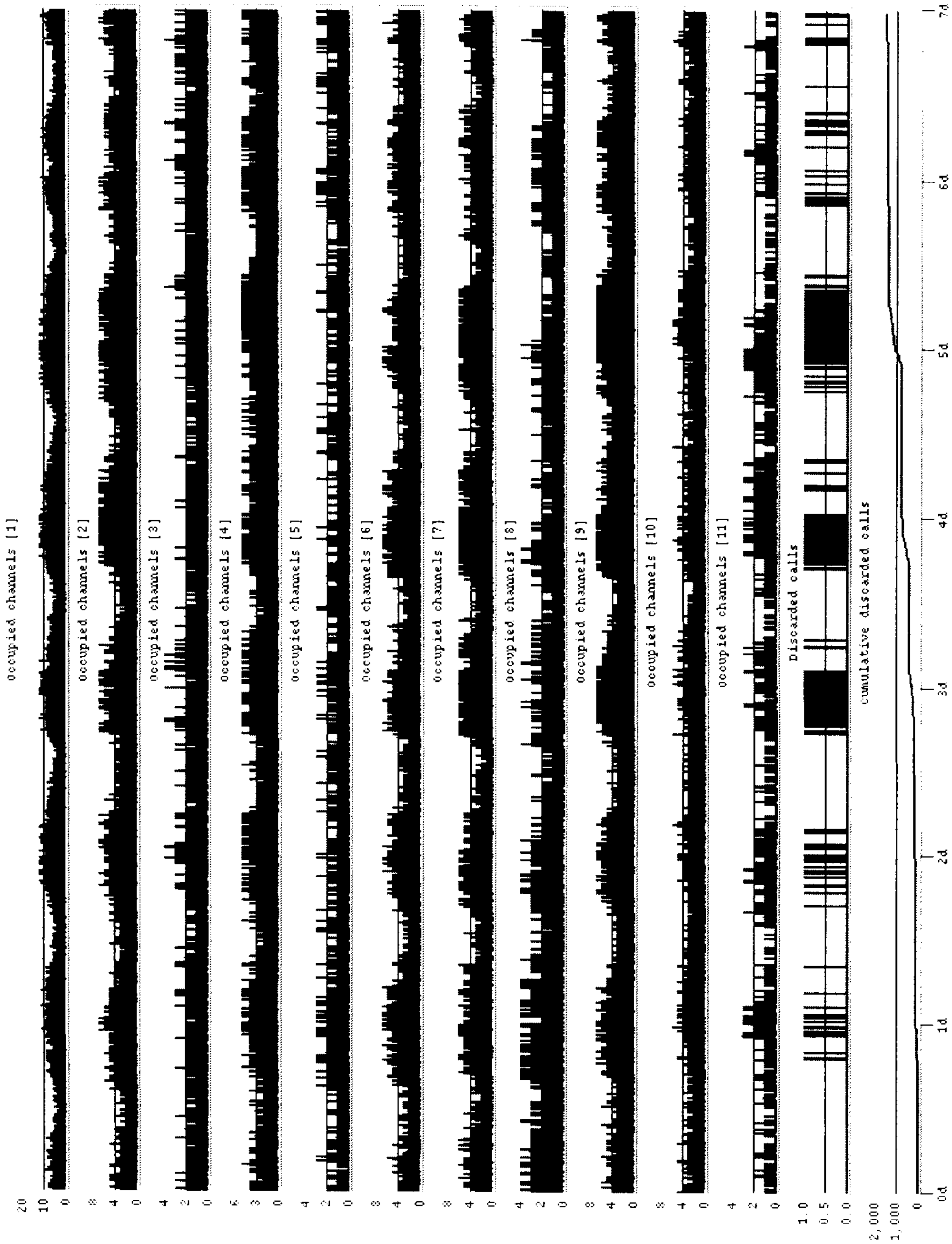

Figure 6.2: OPNET statistics collected during the simulation of network activity in 2003. "Occupied channels" graphs [1] to [11] show the utilization of each cell (number of occupied radio channels). "Discarded calls" graph indicates the time instances when calls are discarded. "Cumulative discarded calls" graph shows the cumulative number of discarded calls.

The "Occupied channels [*i*]" graphs indicate the presence of daily (diurnal) cycles in the activity data. The minimum number of used channels is observed at approximately 2 PM every day, while the maximum utilization is reached between 9 PM and 3 AM. As expected, discarded calls occur during periods of high utilization. The simulation results from the 2002 data, shown in Figure 6.1, indicate that most cells seldom reach their capacity. Cell 11, with a capacity of 3 channels, is an exception and often has all available channels occupied during the busy hours. Cell 10 had all its channels occupied only once during the 2002 sample week. Simulation results for the 2003 data are shown in Figure 6.2. During the busy hours, several cells operate at their full capacities. (Cell capacities are given in Table 4.1. For example, every channel in cells 2, 4, 7, and 9 is occupied during the periods of high utilization. In both 2002 and 2003 sample data, the average number of used channels in each cell is small compared to its capacity. This is to be expected because of the design requirements that the system meets certain grade of service (GoS) criteria during busy hours. The average number of used channels and the average utilization of each cell for both sample weeks are given in Table 6.1. Except for cell 11, the average utilization of every cell increased from 2002 to 2003. The average utilization of cell 4 increased by almost a factor of four.

The OPNET model does not match the behaviour of the deployed system with respect to discarded calls. The trace file (data model) was created from the sample data by considering only the established calls. Discarded calls in the simulation results are due to discrepancies in the sample data. As described in Chapter 3, there are records corresponding to one call that may have multiple values for the timestamp and the duration. Since the trace file (data model) was created by taking the largest value for the call duration, it is possible that the simulation model exhibits larger utilization than the deployed network.

Table 6.1: OPNET model: average number of used channels and average utilization

Cell	2002			2003		
	No. of calls	Average no. of used channels	Average utilization (%)	No. of calls	Average no. of used channels	Average utilization (%)
1	387,640	2.45	20.42	401,212	2.70	22.50
2	132,083	0.80	11.43	202,958	1.33	19.00
3	48,314	0.27	6.75	51,839	0.31	7.75
4	41,741	0.25	5.00	150,033	0.95	19.00
5	24,620	0.14	4.67	32,587	0.19	6.33
6	112,393	0.70	10.00	167,878	1.11	15.86
7	132,662	0.81	13.50	184,634	1.22	20.33
8	36,576	0.22	5.50	60,445	0.37	9.25
9	75,019	0.46	7.67	229,615	1.50	21.43
10	40,863	0.23	3.83	116,491	0.74	12.33
11	56,770	0.33	11.00	26,871	0.18	6.00

Another discrepancy is the existence of records showing overlapping usage of radio channels in the cells. One channel in a given cell can be occupied by only one call at a time. As indicated in Table 6.2, channel 7 in cell 1 is busy with a call lasting from 0:53:57.467 until 0:54:12.287 (= 0:53:57.467 + 14.820). However, there is a record corresponding to a call that originates at 0:54:12.180, occupying the very same channel 7 in cell 1. These records of overlapping usage of channels are due to errors in the recording of the activity data in the deployed network. Therefore, in the simulation scenarios, these calls are established using two channels instead of one. If a second channel is not available, one of the calls is discarded.

Table 6.2: **Overlapping usage of channels.**

Timestamp	Duration (ms)	Cell	Channel
2003-03-10, 0:53:57.467	14,820	1	7
...	...	...	...
2003-03-10, 0:54:12.180	2,640	1	7

The simulation results show 343 and 1,487 discarded calls in 2002 and 2003, respectively. The total number of calls was 477,953 (2002 sample week) and 625,898 (2003 sample week), respectively. The number of discarded calls in OPNET simulations is small compared to the total number of calls and, hence, it does not affect the analysis of the network utilization.

6.2 Analysis of Discarded Calls and Traffic Trends

Grade of service (GoS) is an important parameter when analyzing and provisioning circuit switched networks. It represents the probability of discarded calls for systems without queues or probability of a call being delayed or blocked in systems with queues. Although the data model consists only of established calls, the discarded calls may be used to locate network bottlenecks. They are indicators of network congestion and occur at periods of high network utilization. They may also indicate future network congestion problems if the traffic loads of the busy cells increase.

The busiest cells are most likely to cause call drops. Cell(s) that caused a call to be discarded could not be determined from the collected OPNET simulation statistics. From the graphs shown in Figure 6.1, it is only possible to identify the cells that operate near the full capacities (given in Table 4.1). For example, as shown in Figure 6.1, cell 11 and, to a smaller extent, cells 2 and 7 have every channel occupied during busy hours in

2002 when most of the discarded calls occur. Similarly, from the simulation results for the 2003 sample week shown in Figure 6.2, cells 2, 4, 7, and 9 may be identified as the busiest cells. We varied the capacities of cells 2, 7, and 11 (2002 data) and cells 2, 4, 7, and 9 (2003 data) as indicated in Table 6.3, while maintaining capacities of the remaining cells as shown in Table 4.1. The number of discarded calls decreased significantly.

Table 6.3: OPNET simulation results for various cell capacities.

<table>
<tr><th>Sample data</th><th>Cell no.</th><th>No. of channels</th><th>No. of discarded calls</th></tr>
<tr><td rowspan="6">2002</td><td colspan="2">Original capacities</td><td>343</td></tr>
<tr><td>11</td><td>3 + 1</td><td>113</td></tr>
<tr><td>11</td><td>3 + 2</td><td>91</td></tr>
<tr><td>11</td><td>3 + 1</td><td rowspan="3">58</td></tr>
<tr><td>7</td><td>6 + 1</td></tr>
<tr><td>2</td><td>7 + 1</td></tr>
<tr><td rowspan="7">2003</td><td colspan="2">Original capacities</td><td>1,487</td></tr>
<tr><td>9</td><td>7 + 1</td><td rowspan="2">863</td></tr>
<tr><td>7</td><td>6 + 1</td></tr>
<tr><td>9</td><td>7 + 1</td><td rowspan="4">435</td></tr>
<tr><td>7</td><td>6 + 1</td></tr>
<tr><td>4</td><td>5 + 1</td></tr>
<tr><td>2</td><td>7 + 1</td></tr>
</table>

This implies that the originally identified busiest cells indeed contributed to most of the discarded calls.

Unlike the OPNET model, WarnSim [2], [43] is designed to collect statistics of blocked and discarded calls for each cell. Hence, it is possible to determine the number of discarded calls in each cell and to identify the cells that experience most of the discarded calls. The number of discarded calls and the call blocking probability for each cell for the two sample weeks are shown in Table 6.4 [2].

Table 6.4: WarnSim: simulation results of discarded calls and call blocking probabilities.

	2002		2003	
Cell	Discarded calls	Call blocking probability (%)	Discarded calls	Call blocking probability (%)
1	7	0.00	30	0.01
2	15	0.01	211	0.10
3	0	0.00	2	0.00
4	2	0.00	290	0.19
5	5	0.02	9	0.03
6	7	0.01	41	0.02
7	44	0.03	513	0.28
8	0	0.00	4	0.01
9	1	0.00	301	0.13
10	0	0.00	5	0.00
11	237	0.42	38	0.14

Call blocking probability in a cell is calculated as the number of discarded calls divided by the total number of calls in the cell. In the 2002 sample week, the majority of discarded calls occurred in cell 11, followed by cells 2 and 7. In 2003, cells 2, 4, 7, and 9 contributed to most of the discarded calls. This agrees with the OPNET results shown in Table 6.3

The differences in network traffic loads in 2002 and 2003 were shown in Table 6.1 and in Figure 6.1 and Figure 6.2. The number of calls increased by ~30%. This increase was not uniform among the cells. Several cells (4, 9, and 10) experienced a larger increase in the traffic load. If the same trend of non-uniform increase of traffic load persists, cells 4, 9, and 10 may experience queued and discarded calls. As shown, queued and discarded calls may be avoided by increasing the capacities of these cells.

WarnSim simulation results indicate that, except for cell 11, the number of discarded calls increased from 2002 to 2003, [2]. Call blocking probability also increased in most cells. The increase in discarded calls and call blocking probability is significant for cells 2, 4, 7, and 9, which implies decreased grade of service (GoS) in these cells.

6.3 Statistical Modeling

The main goal of statistical analysis is to extract user behaviour patterns that are contained in traffic traces excerpt from E-Comm's records detail file. These patterns could be useful for future prediction of traffic in the system and for better understanding and control of the network in whole. The data from traffic traces could be treated as time series.

Two main approaches that could be employed while analysing time series are the time domain and the frequency domain approach. The first of them represents time series as function of time while second treats time series as spectral evaluation of wavelets or Fourier expressions.

The time series plots of the busiest hour for call holding time over a period of three consecutive years 2001, 2002, and 2003 are shown in Figure 6.3, Figure 6.4, and Figure 6.5. The call inter-arrival time series plots are given in Figure 6.6, Figure 6.7, and Figure 6.8 for 2001, 2002, and 2003 years respectively. The busiest hours in System 1 were 15:00-16:00 on 01.11.2001, 4:00-5:00 on 01.03.2002 and 22:00-23:00 on

26.03.2003.The y- axes represent call holding times or call inter-arrival times respectively. The values on x-axis represent sample number in the series. These plots allow insight in the data sets and detection of outliers if there are any. Such outlier detection and removal is extremely important for identifying irregularities in the data acquisition system and for cleaning data sets for modeling. These plots also could be useful to discover a variety of structures as is trends or irregular components or (quasi-) periodic components (seasonal component or some other with obvious cycles).

By visual inspection of call holding time series plots, we could notice a slight increase in call holding time when we compare the busiest hour from 2001 with the busiest hours from 2002 and 2003. For example, we could notice that there are more calls that last more than 25 seconds in following years than in 2001 year. Furthermore, the number of calls and the average duration of calls in the busiest hour increase from 2001 to 2002 and to 2003.

The data traces do not reveal any obvious visible tendencies, which could imply non-stationarity. Stationary processes have the property that autocorrelation, the mean, and the variance do not change over time and do not have periodic fluctuations. The testing results for wide sense or second order stationarity for call inter-arrival and call holding times are given in Sections 6.4 and 6.5.

Time series plots of call inter-arrival time of three busiest intervals of these years reveal that call inter-arrival time decreases when compare 2001, 2002 and 2003 years. For example, if we consider time distance between two calls of six seconds there are more cases in the sample of the year 2001 than in the following years. The average values of call inter-arrival times confirm conjecture about the decreasing trend.

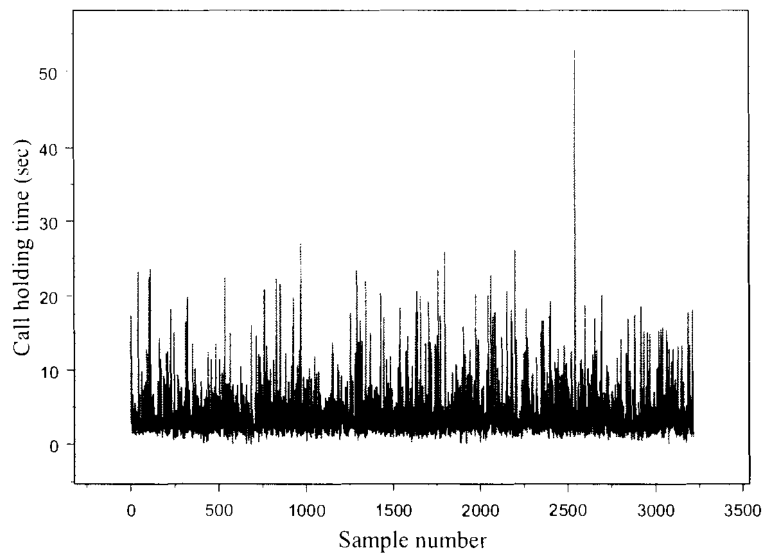

Figure 6.3: Call holding time during busy period from 15h to 16h on 01.11.2001.

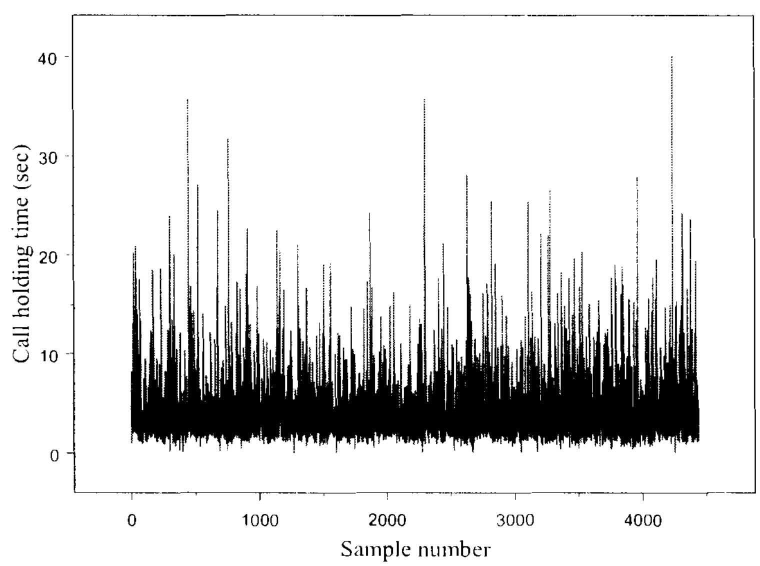

Figure 6.4: Call holding time during busy period from 4h to 5h on 01.03.2002.

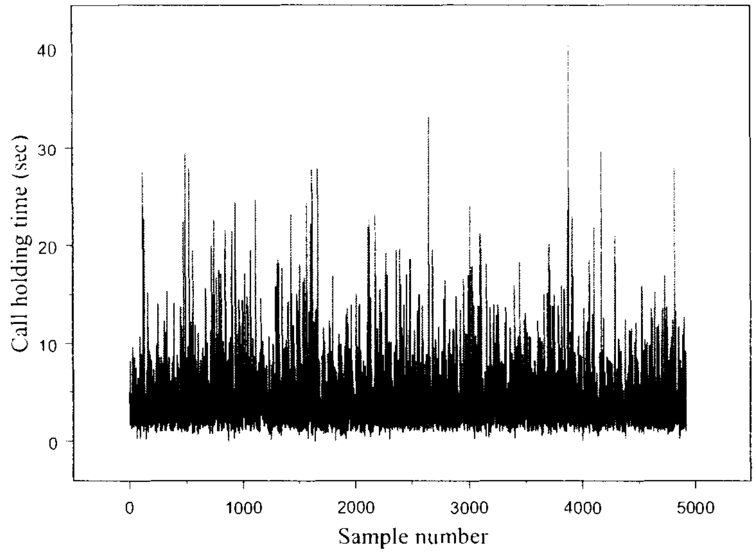

Figure 6.5: Call holding time during busy period from 22h to 23h on 26.03.2003.

This implies (taking in consideration increasing of call holding time) that last two years were busier in the busy hour period than 2001. This finding was expected because of larger number of network users and it agrees with OPNET simulation results [1]. For example, the analysis of utilization of E-Comm network over period of two sample weeks: February 1-8, 2002 and March 20-26, 2003 showed that number of calls increased by nearly 60% between February 2002 and March 2003 [1], so that we could also expected increasing number of calls during busy hours in these two years. These graphs do not provide an opportunity to make conclusion regarding detail difference between three years and further analysis is necessary.

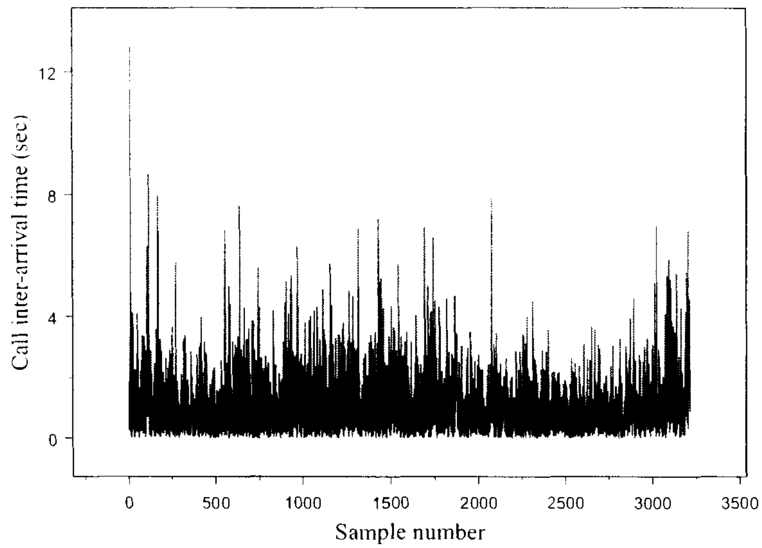

Figure 6.6: Call inter-arrival time during busy period from 15h to 16h on 01.11.2001.

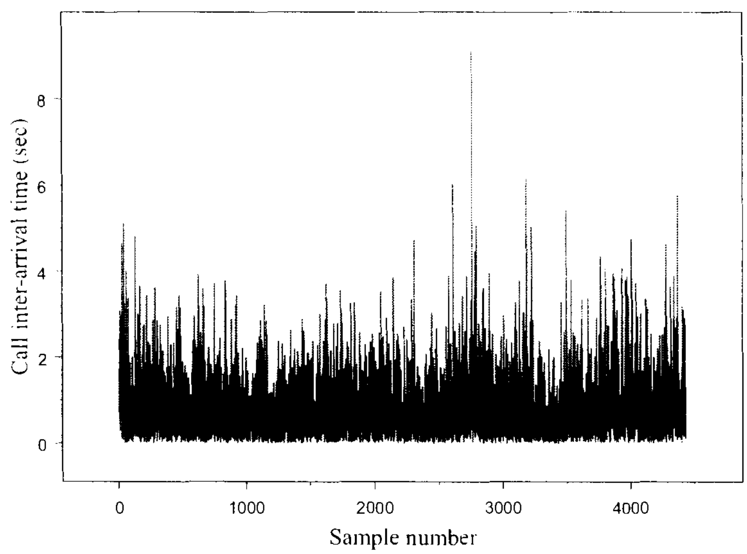

Figure 6.7: Call inter-arrival time during busy period from 4h to 5h on 01.03.2002.

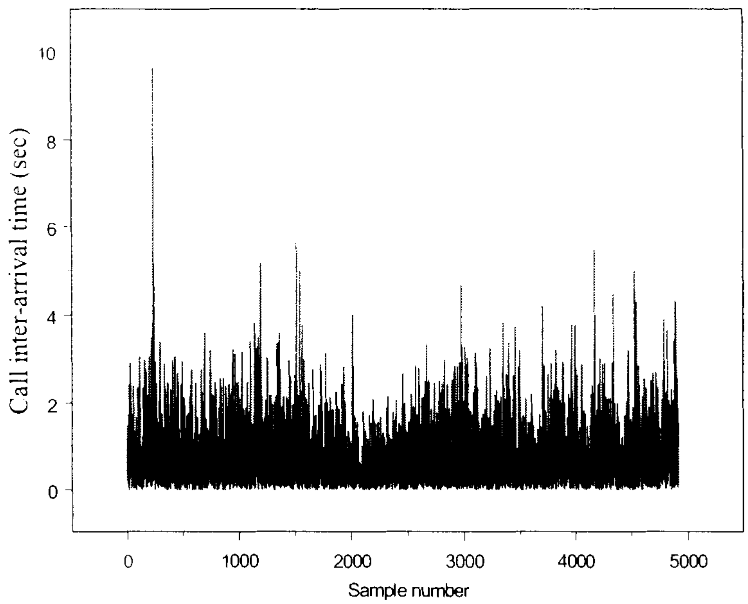

Figure 6.8: Call inter-arrival time during busy period from 22h to 23h on 26.03.2003.

The most standard statistical test depends on randomness (no time dependences in time series). Hence, it was necessary to check validity of randomness assumption in our data samples. The validity of the test conclusions is directly connected to the validity of the randomness assumption. For example, most widely-used formulae for determining standard deviation of sample mean ($S_{\bar{Y}} = S / \sqrt{N}$) is not valid if randomness assumption doesn't hold and distributions plots are not meaningful when data are not random [52].

Commonly used tool for testing randomness are correlograms (autocorrelation plots that show autocorrelation coefficients plotted versus lag) and lag plots. A lag is a fixed time displacement [52].

Autocorrelation plots ascertain randomness by calculating autocorrelation coefficient for data set at varying time lags. These coefficients should be close to zero for all lags if data are random, otherwise one or more of them will significantly differ of zero. Following is formulae for computing autocorrelation coefficients:

$$R_h = C_h / C_0, \tag{11}$$

where C_h is the autocovariance function

$$C_h = \left(1/N\right) \sum_{i=1}^{N-h} \left(Y_i - \bar{Y}\right)\left(Y_{i+h} - \bar{Y}\right), \tag{12}$$

and C_0 is the variance function

$$C_0 = \frac{\sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2}{N}. \tag{13}$$

Autocorrelation coefficients are always between -1 and 1 and they enable detection of presence of serial correlation (consecutive values in time series correlated one with another). This presence could leads to increased spectral power at lower frequencies (redness) what could be seen on spectral density graph of our data sets. In addition, serial correlation and trends could reduce the effective number of degrees of freedom in time series.

Autocorrelation plots could be useful for prediction models as is autoregressive moving average time series models. It is important to mention that confidence bands 95% and 99% are calculated on different way depends whether autocorrelation plot is used to test for randomness or in identification prediction models if data are not random. Since, in this moment, we are particularly interested for time dependences in data set, the next formulae $\pm\frac{z_{1}-\alpha/2}{\sqrt{N}}$ is applied for calculating confidence interval, where N is the sample size, z is the percent point function of the standard normal distribution and α is significance level. Obviously the confidence bands depends of sample size and their width will be constant for all lags in this case.

Autocorrelation coefficients plots versus lag for call holding time during the busiest period in three consecutive years with 95% and 99% confidence interval are shown in Figure 6.9, Figure 6.10, and Figure 6.11.

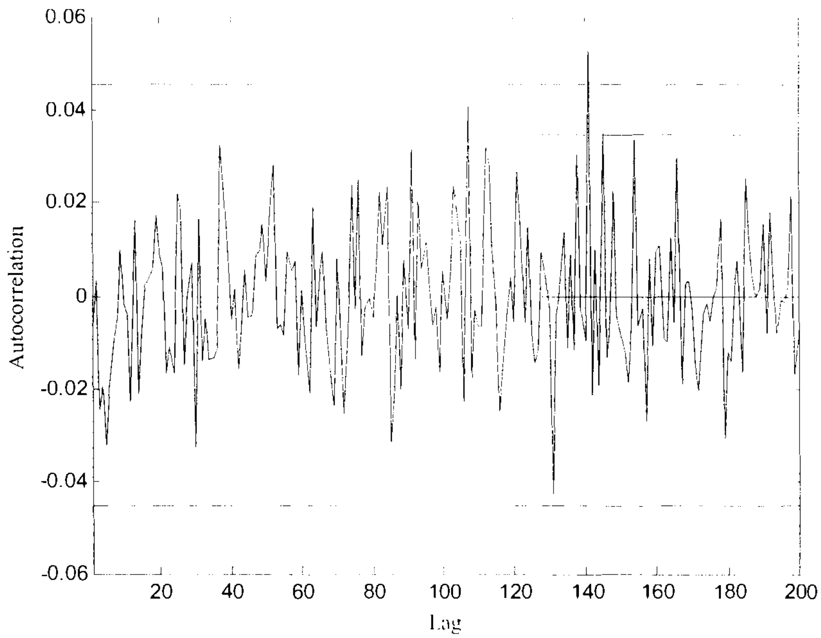

Figure 6.9: Call holding time autocorrelation plot during busy period from 15h to 16h on 01.11.2001.

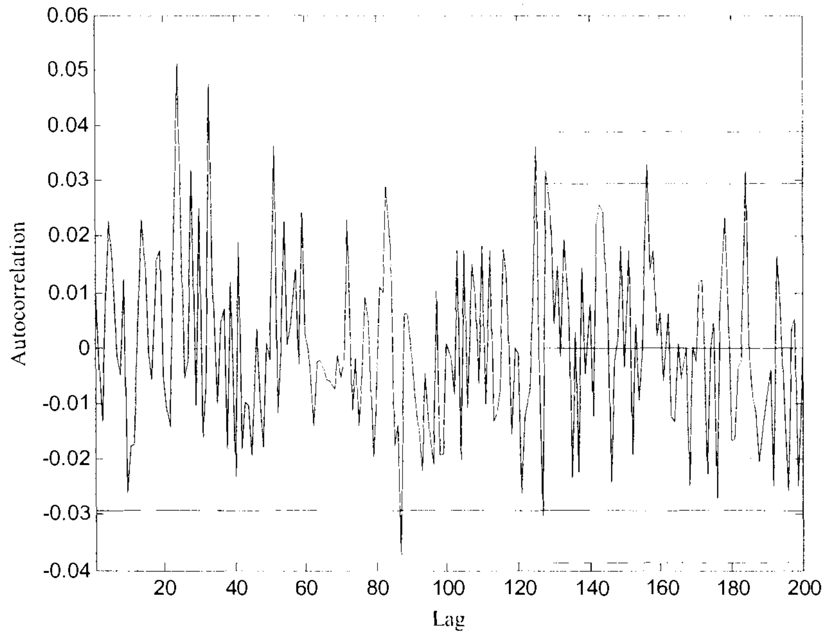

Figure 6.10: Call holding time autocorrelation plot during busy period from 4h to 5h on 01.03.2002.

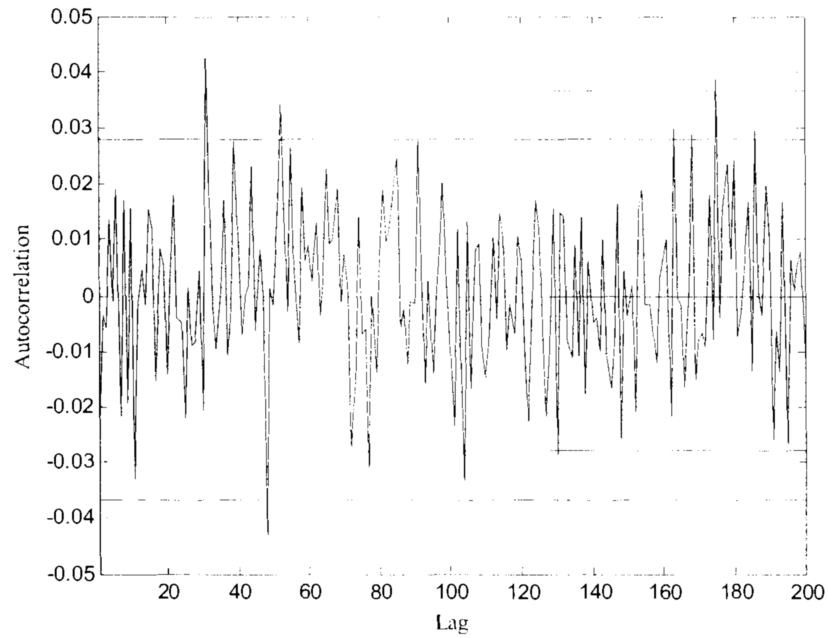

Figure 6.11: Call holding time autocorrelation plot during busy period from 22h to 23h on 26.03.2003.

These plots reveal that there is no significant correlation for any time-lag separations. All values except a few of them reside within the 95 % confidence interval around zero and for these confidence limits we might expect one out of twenty lags out of bounds to be statistically significant. In addition, there is no apparent pattern. There is no conclusion that data from 2002 and 2003 exhibit more correlation although a few values more fall outside confidence bands. Autocorrelation function peaks appear for separation in number of samples that correspond to two not mandatory consecutive peaks in time series plots. Note that, no any of our autocorrelation plots show results for lag 0. They are moved in right since for lag0, autocorrelation coefficient is always 1 by definition.

Figure 6.12, Figure 6.13, and Figure 6.14 depict autocorrelation plots for inter-arrival time of busiest hour in 2001, 2002 and 2003 year respectively with 95% and 99% confidence interval. These plots illustrate that call inter-arrival time exhibit significant correlation for all years approximately up to lag 70. After lag 70 only a few values lay outside 95% confidence interval around zero.

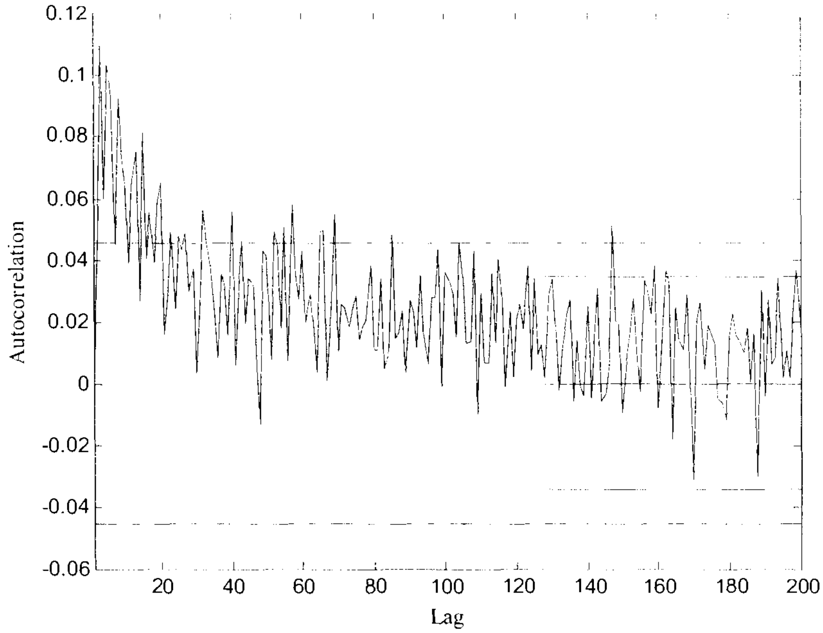

Figure 6.12: Call inter-arrival time autocorrelation plot during busy period from 15h to 16h on 01.11.2001.

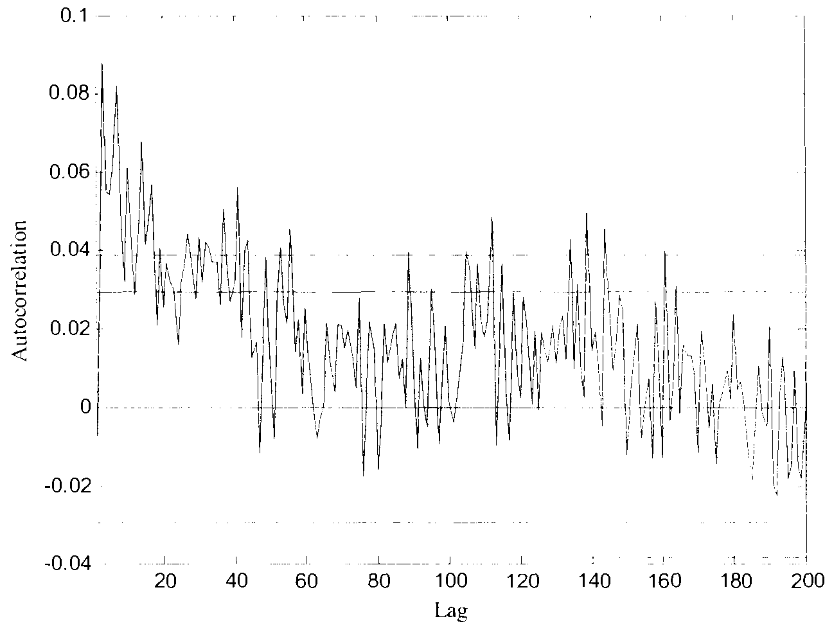

Figure 6.13: Call inter-arrival time autocorrelation plot during busy period from 4h to 5h on 01.03.2002.

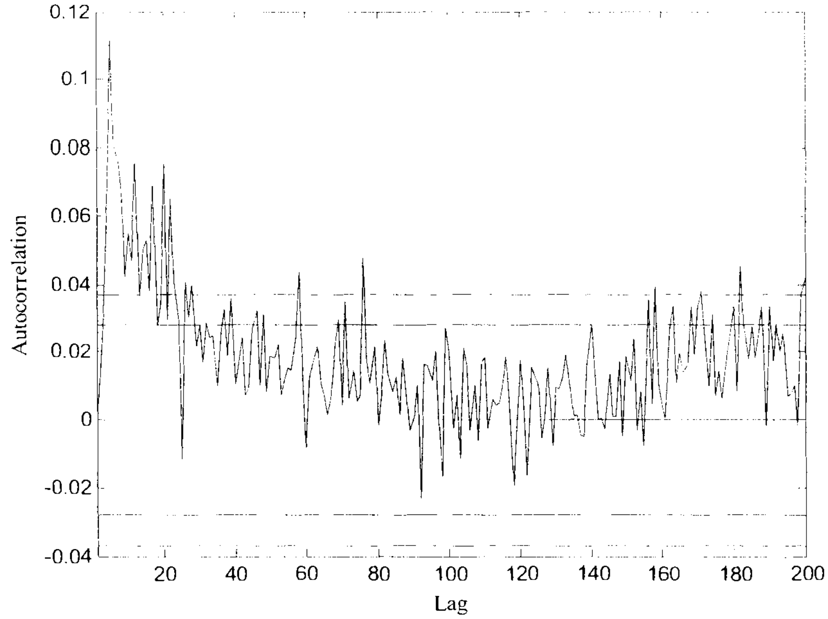

Figure 6.14: Call inter-arrival time autocorrelation plot during busy period from 22h to 23h on 26.03.2003.

The second routine tool for checking randomness is lag plot. The lag plot is essentially a scatter plot with two variables properly lagged [52]. Random data not exhibit any identifiable structure in lag plot while non-random data exhibits some pattern apropos non-random structure. The lags plots for lag 1 are generated since it is most commonly used lag, in addition by a few other plots for arbitrary lags to ascertain that plots have similar appearances.

Plots of lag 1 of call holding time respectively for busiest hour in 2001, 2002 and 2003 year are shown in Figure 6.15, Figure 6.16, and Figure 6.17. Based on these plots we could make following conclusions: the data do not exhibit autocorrelation, the data do not contain outliers and the data are random.

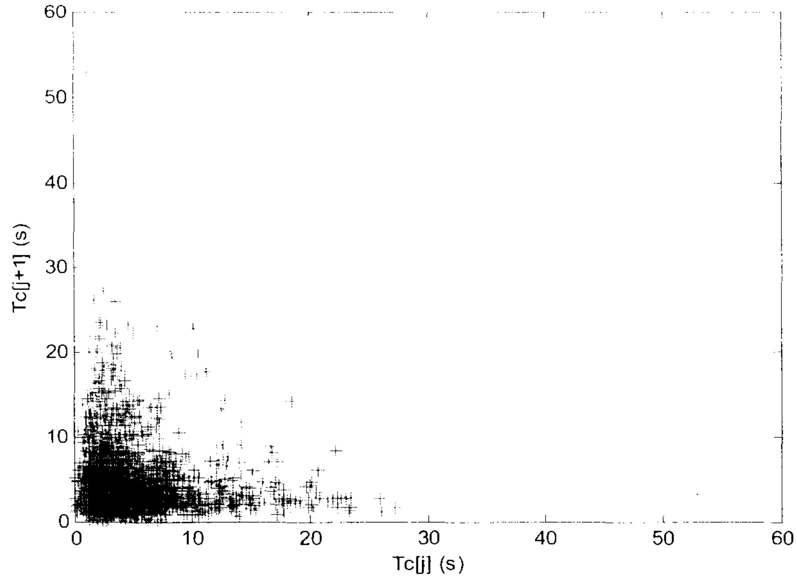

Figure 6.15: Call holding time Tc lag plot during busy period from 15h to 16h on 01.11.2001.

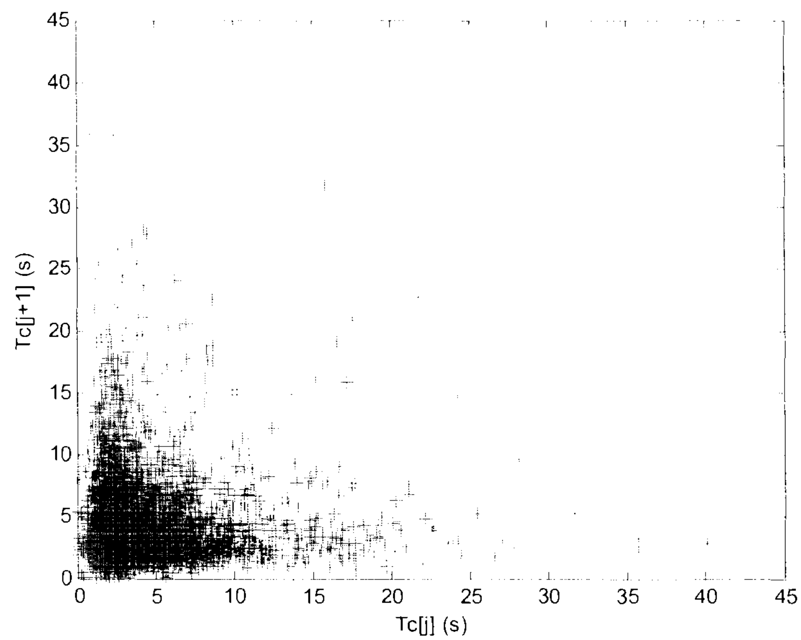

Figure 6.16: Call holding time Tc lag plot during busy period from 4h to 5h on 01.03.2002.

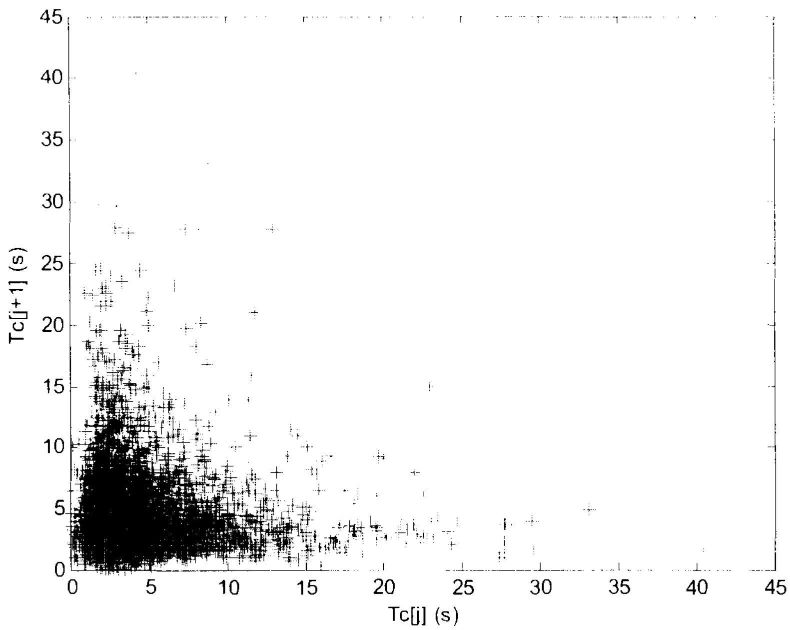

Figure 6.17: Call holding time Tc lag plot during busy period from 22h to 23h on 26.03.2003.

Figure 6.18, Figure 6.19, and Figure 6.20 represent lag plots for call inter-arrival time for the busiest hour in 2001, 2002 and 2003 year respectively. Absence of visible structure implies that data are random since we cannot conclude what is going to be next value of call inter-arrival time based on current value. It could be virtually anything between 0 and 10 seconds. This non-association ascertains randomness. In addition, it is obvious that there are no any significant outliers or clusters in the data set

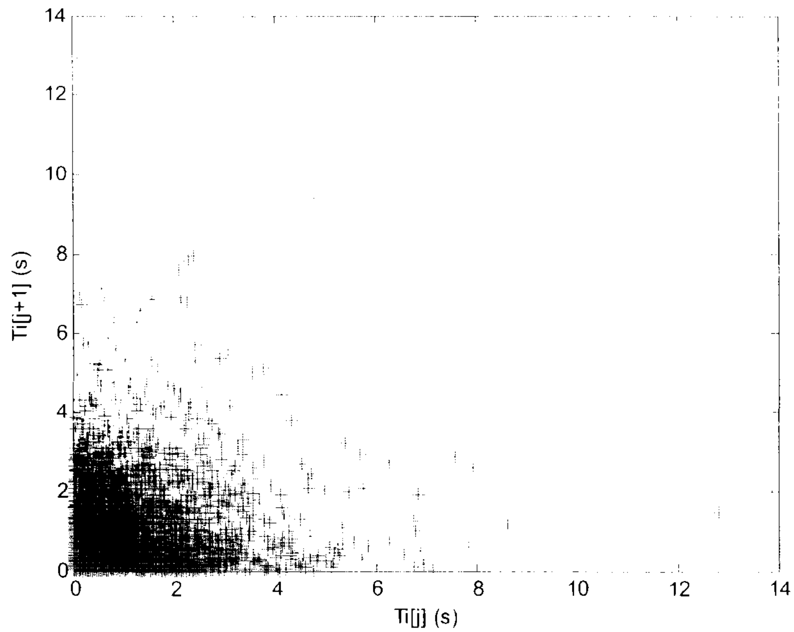

Figure 6.18: Call inter-arrival time Ti lag plot during busy period from 15h to 16h on 01.11.2001.

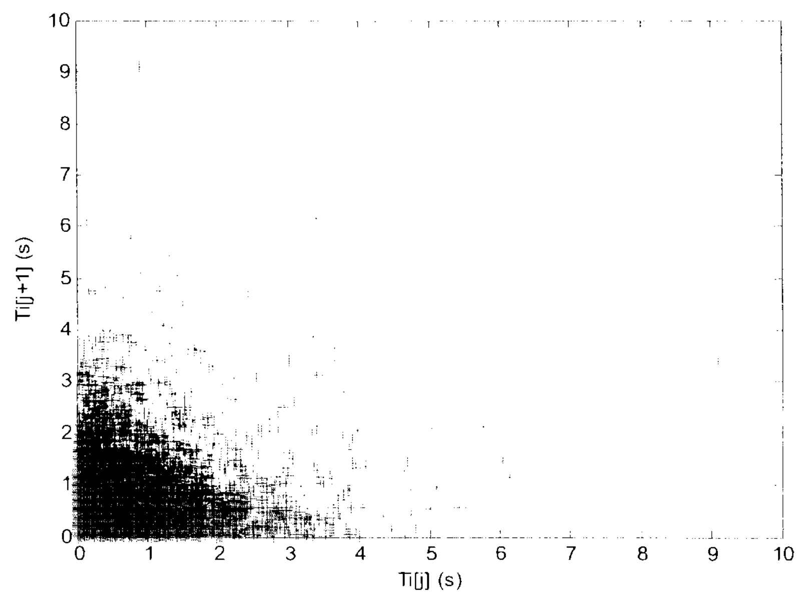

Figure 6.19: Call inter-arrival time Ti lag plot during busy period from 4h to 5h on 01.03.2002.

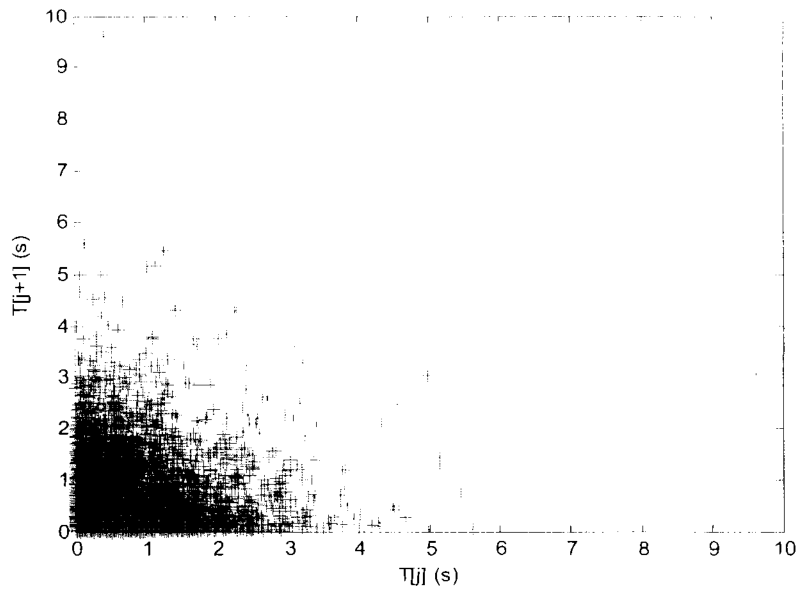

Figure 6.20: Call inter-arrival time Ti lag plot during busy period from 22h to 23h on 26.03.2003.

In order to discover possible clusters in E-comm data sets more detailed clustering analysis was employed [35], [36]. K-means algorithm and AutoClass was employed on the network log data collected from March 1. 2003 to May 31. 2003. AutoClass discovered 24 populated clusters of talk groups as the best classification, with average cluster size of 25.7, herewith that the first three clusters represent over 90% of all talk groups. K-means algorithm confirmed that clustering with three clusters is the most useful for prediction of future traffic in the E-comm network [35], [36]. These lags plots were created in order to confirm randomness assumption and to detect possible outliers, if any.

Since data do not exhibit significant non-randomness, statistical tests are employed to discover if any of well-known distributions underlie given sets of data. We analyze traces of call holding and call inter-arrival times. For each trace, we determine the best fitting distribution using the K-S test and examine the autocorrelation function.

We also test the traces for long-range dependence by performing wavelet-based estimation of the Hurst parameter and by testing the time constancy of the scaling exponent α.

The figures included in Sections 6.4 and 6.5 illustrate graphical results obtained from the 2003 busiest hour (between 22:00 and 23:00 on 26 March). The results from other hourly traces are similar.

6.4 Call Inter-Arrival Times

We consider the following fourteen candidate distributions: exponential, Weibull, gamma, normal, lognormal, logistic, log-logistic, Nakagami, Rayleigh, Rician, t-location scale, Birnbaum-Saunders, extreme value, and inverse Gaussian. The parameters of the distributions are calculated by performing Maximum Likelihood Estimation using the hourly traces of call inter-arrival times. For each trace, we plot the distribution of the trace (histogram) and the probability density function for each distribution. In addition, CDF, Quantile (inverse CDF), probability plots, survivor function and cumulative hazard function with parameters estimated from underlying data samples were produced. Examples of the plots are shown in Figure 6.21.

It can be concluded that exponential, lognormal, Weibull, and gamma distributions fit the data better than the others. Quantitative results of the two-sided K-S goodness-of-fit test for the 2001, 2002, and 2003 datasets are shown in Table 6.5,

Table 6.6 and Table 6.7 respectively. The test is performed on the entire traces. The significance level α is 0.1. The null hypothesis (distribution fits the data) is rejected if the test statistics k is greater than the critical value. In that case, the test returns h=1. Otherwise, the test returns h=0, which implies that the distribution fits the data

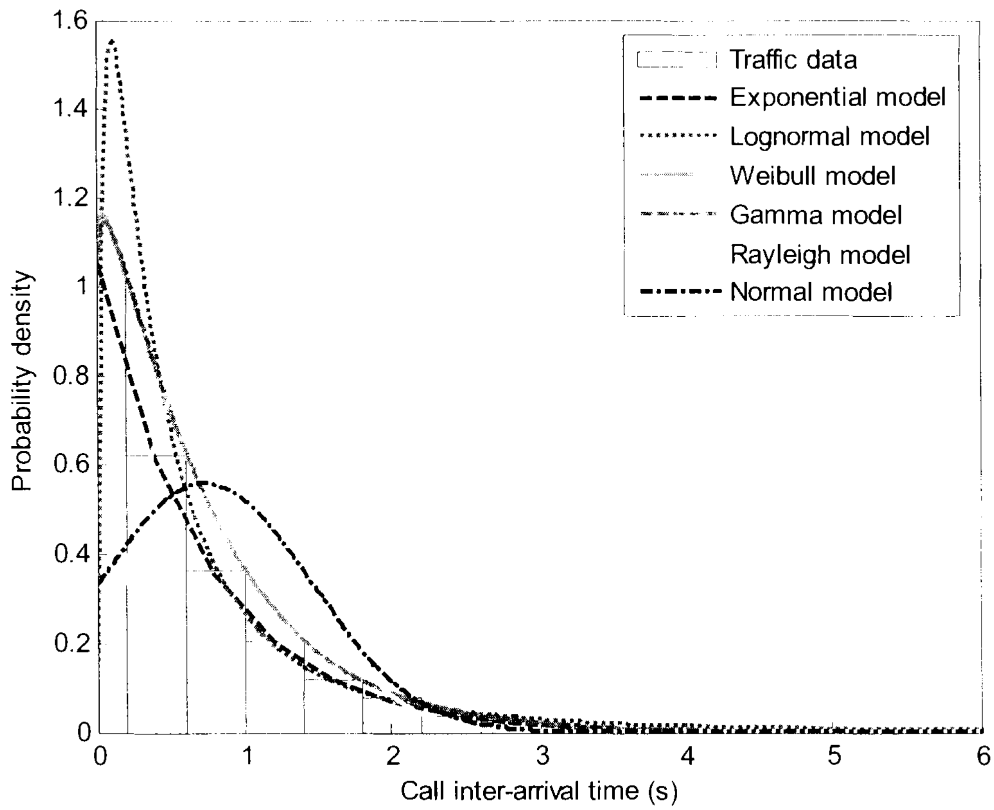

Figure 6.21: Call inter-arrival times distributions.

Both Weibull and gamma distributions fit the call inter-arrival times in E-Comm network, as indicated in Table 6.5,

Table 6.6, and Table 6.7. The test accepts the null hypotheses with a significance level of 0.1 for most hourly traces. Weibull distribution fails the test for two traces in 2001. Gamma fails the test for three traces in 2001. We repeat the K-S test for those traces and the corresponding distributions for various values of α. The distributions may be accepted for smaller significance levels (between 0.01 and 0.1). In contrast, test results show that the lognormal distribution cannot be accepted as a suitable model because it did not pass the test for any trace and for any α. The exponential distribution also fails the test for the majority of traces. Nevertheless, its p-values are significantly higher than those of the lognormal distribution, which indicates a better fit.

Figure 6.22 shows the cumulative distribution function of the call inter-arrival times from the 2003 busiest hour and the exponential, Weibull, and gamma distributions. The four curves almost overlap, indicating a good fit.

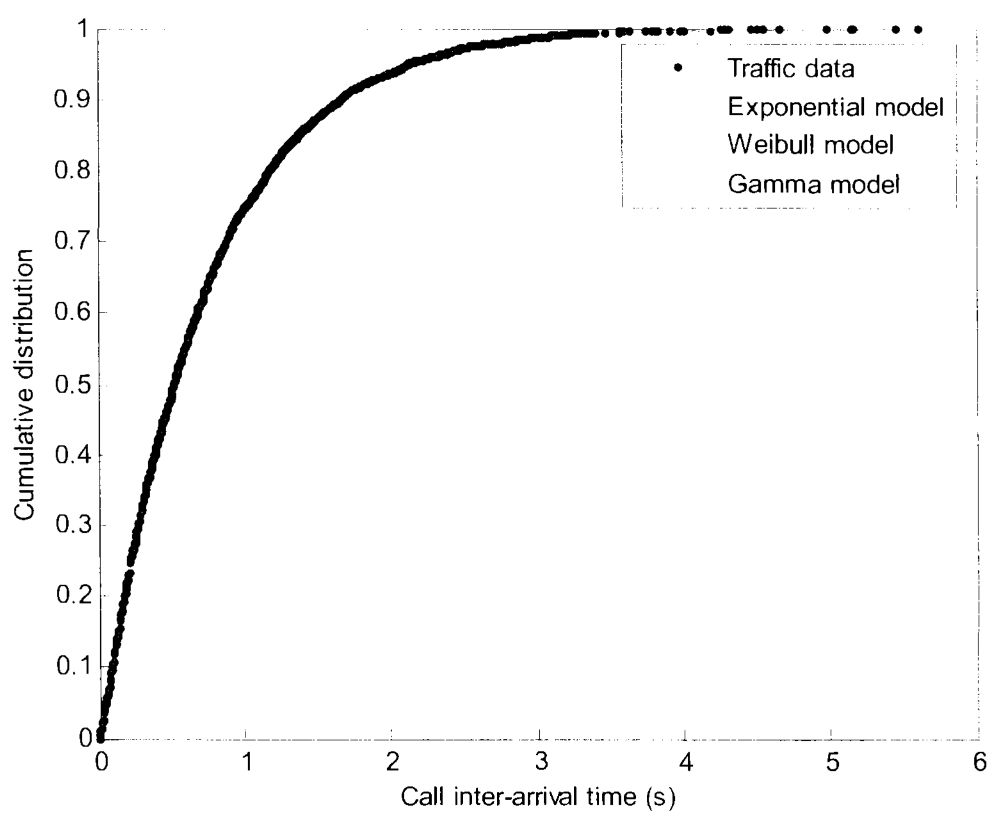

Figure 6.22: Cumulative distribution function of the call inter-arrival times and comparison with exponential, Weibull, lognormal, and gamma distributions.

Table 6.5: K-S test results for the hourly traces of call inter-arrival times from the 2001 busy hours.

Distribution	Param.	02.11.2001 15:00–16:00	01.11.2001 00:00–01:00	02.11.2001 16:00–17:00	01.11.2001 19:00–20:00	02.11.2001 20:00–21:00
Exponential	h	1	1	0	1	1
	p	0.0384	1.45E-04	0.5416	0.0122	0.0135
	k	0.0247	0.0369	0.0131	0.0277	0.0259
Weibull	h	0	1	0	0	1
	p	0.3036	0.0409	0.4994	0.1574	0.0837
	k	0.0171	0.0236	0.0136	0.0195	0.0206
Gamma	h	0	1	0	1	1
	p	0.3833	0.0062	0.3916	0.0644	0.0953
	k	0.0159	0.0287	0.0148	0.0227	0.0202
Lognormal	h	1	1	1	1	1
	p	1.6520E-19	7.0722E-17	1.5936E-22	2.3731E-16	1.3828E-22
	k	0.0769	0.0713	0.0853	0.0743	0.0658

Table 6.6: K-S test results for the hourly traces of call inter-arrival times from the 2002 busy hours.

Distribution	Param.	01.03.2002 04:00-05:00	01.03.2002 22:00–23:00	01.03.2002 23:00–24:00	01.03.2002 00:00–01:00	02.03.2002 00:00–01:00
Exponential	h	1	1	1	1	1
	p	0.0105	6.58E-05	0.0089	0.0304	9.53E-04
	k	0.0243	0.0351	0.025	0.0229	0.0311
Weibull	h	0	0	0	0	0
	p	0.3865	0.6608	0.3384	0.8846	0.5918
	k	0.0136	0.0113	0.0143	0.0093	0.0123
Gamma	h	0	0	0	0	0
	p	0.2007	0.6704	0.5017	0.9135	0.4921
	k	0.0161	0.0112	0.0126	0.0089	0.0133
Lognormal	h	1	1	1	1	1
	p	3.0769E-19	2.4339E-16	9.9945E-16	1.7524E-18	4.0452E-19
	k	0.0698	0.0651	0.0648	0.0723	0.0739

Table 6.7: **K-S test results for the hourly traces of call inter-arrival times from the 2003 busy hours.**

Distribution	Param.	26.03.2003, 22:00–23:00	25.03.2003, 23:00–24:00	26.03.2003, 23:00–24:00	29.03.2003, 02:00–3:00	29.03.2003, 01:00–02:00
Exponential	h	1	1	0	1	1
	p	0.0027	0.0469	0.4049	0.0316	0.1101
	k	0.0283	0.0214	0.0137	0.0205	0.0185
Weibull	h	0	0	0	0	0
	p	0.4885	0.4662	0.2065	0.286	0.0837
	k	0.013	0.0133	0.0164	0.014	0.0159
Gamma	h	0	0	0	0	0
	p	0.3956	0.3458	0.127	0.145	0.1672
	k	0.0139	0.0146	0.0181	0.0163	0.0171
Lognormal	h	1	1	1	1	1
	p	1.0147E-20	4.717E-15	2.97E-16	3.2665E-23	4.8505E-21
	k	0.0689	0.0629	0.0657	0.0795	0.0761

The autocorrelation function of the call inter-arrival times from the 2003 busiest hour is shown in Figure 6.23. The horizontal lines show the 95% (dashed line) and 99% (dash-dotted line) confidence intervals, respectively. The majority of the autocorrelation coefficients for lags smaller than 60 are outside the confidence intervals. This indicates presence of non-negligible correlation among call inter-arrival times.

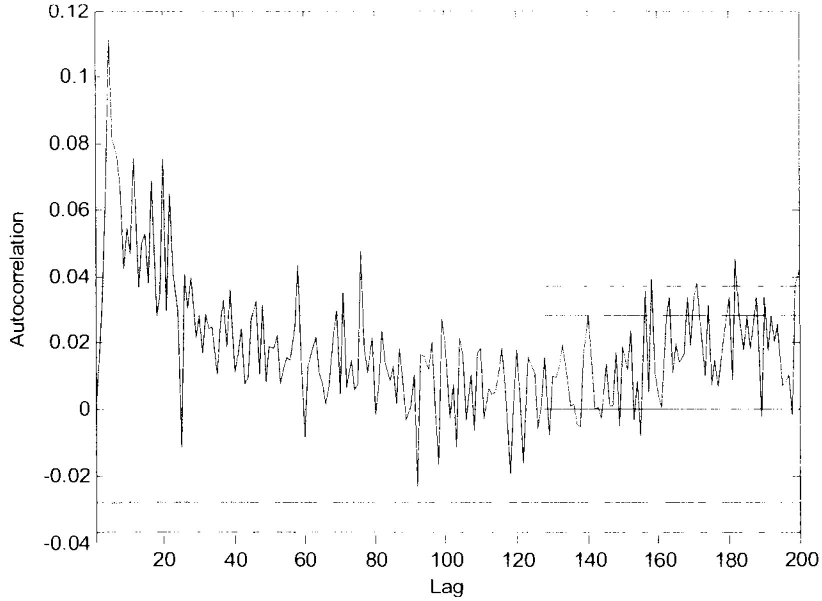

Figure 6.23: Call inter-arrival times autocorrelation plot (up to lag 200) with 95% and 99% confidence intervals.

We test the traces of call inter-arrival times for long-range dependence by performing wavelet-based estimation of the Hurst parameter. An example of a logscale diagram is shown in Figure 6.24. It exhibits a linear relationship between $\log_2 E\{d(j,k)^2\}$ and j with a positive slope in the range of octaves [4–9], which indicates LRD. Estimates of H for all hourly traces are shown in Table 6.8. We also test the time constancy of the scaling exponent α by dividing each trace into m sub-traces, $m \in \{3, 4, 5, 6, 7, 8, 10\}$. All traces pass the test for more than 50% of m's, which indicates that α can be considered constant across the traces and the estimates of H reported in Table 6.8 are reliable. For all traces $H \in (0.5, 1)$, indicating that call inter-arrival times exhibit long-range dependence.

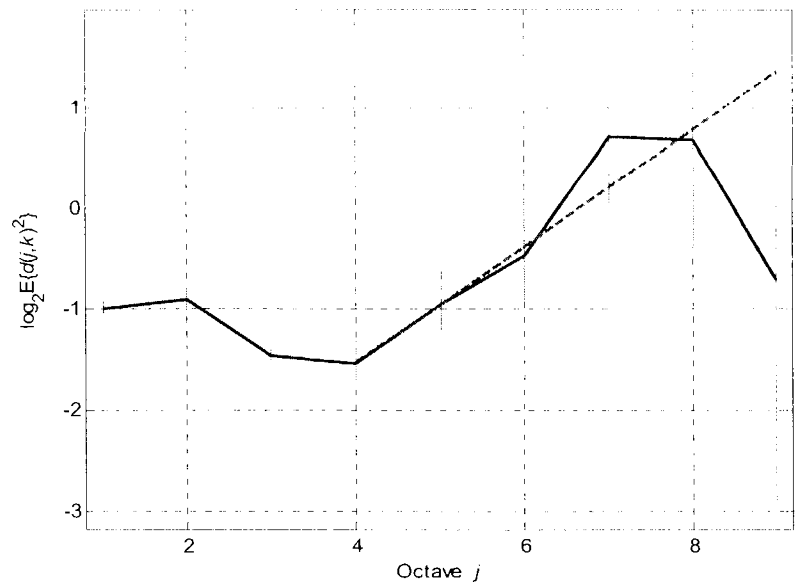

Figure 6.24: Logscale diagram of the call inter-arrival times. Dashed line is the linear regression line with slope α. Vertical lines represent the 95% confidence intervals around the estimates of $\log_2 E\{d(j,k)^2\}$.

Table 6.8: Estimates of H for the hourly traces of call inter-arrival times.

2001		2002		2003	
Day/hour	H	Day/hour	H	Day/hour	H
02.11.2001 15:00–16:00	0.907	01.03.2002 04:00–05:00	0.679	26.03.2003 22:00–23:00	0.788
01.11.2001 00:00–01:00	0.802	01.03.2002 22:00–23:00	0.757	25.03.2003 23:00–24:00	0.832
02.11.2001 16:00–17:00	0.770	01.03.2002 23:00–24:00	0.780	26.03.2003 23:00–24:00	0.699
01.11.2001 19:00–20:00	0.774	01.03.2002 00:00–01:00	0.741	29.03.2003 02:00–03:00	0.696
02.11.2001 20:00–21:00	0.663	02.03.2002 00:00–01:00	0.747	29.03.2003 01:00–02:00	0.705

6.5 Call Holding Times

We compare the distribution of the call holding times with the distributions that were considered for the call inter-arrival times. The probability density function of the call holding times and of the several best fitting distributions are shown in Figure 6.25. The K-S goodness-of-fit test is performed for the lognormal, exponential, Weibull, and gamma distributions. None of the distributions passes the test when the entire traces are tested with significance levels 0.1 and 0.01. Therefore, we perform the test on ten randomly chosen sub-traces of length 1,000 extracted from each trace, with a significance level α of 0.01. Only lognormal distribution passes the test for very few sub-traces. When sub-traces of length 500 are tested with the same significance level, the lognormal distribution exhibits the best fit. It passes the K-S test for almost all 500-sample sub-traces of all hourly traces. The test rejects the null hypothesis when those sub-traces are compared with the other three candidate distributions: exponential, Weibull, and gamma. Figure 6.26 shows the cumulative distribution function of the call holding times from the 2003 busiest hour. It also shows the cumulative distribution function of the exponential, lognormal, Weibull, and gamma distributions. Lognormal distribution fits best the call holding times, as shown in Figure 6.25 and Figure 6.26.

The autocorrelation function of the call holding times from the busiest hour in the 2003 dataset is shown in Figure 6.27. We observe that there are no significant correlations for non-zero lags because all but a few autocorrelation coefficients are within the 95% and 99% confidence intervals.

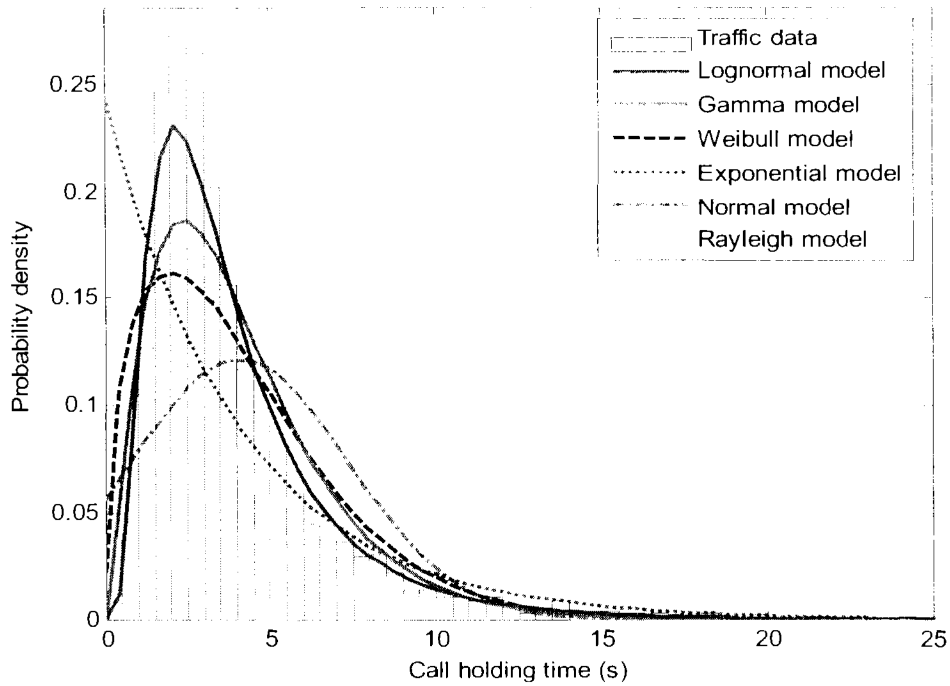

Figure 6.25: Call holding times distributions.

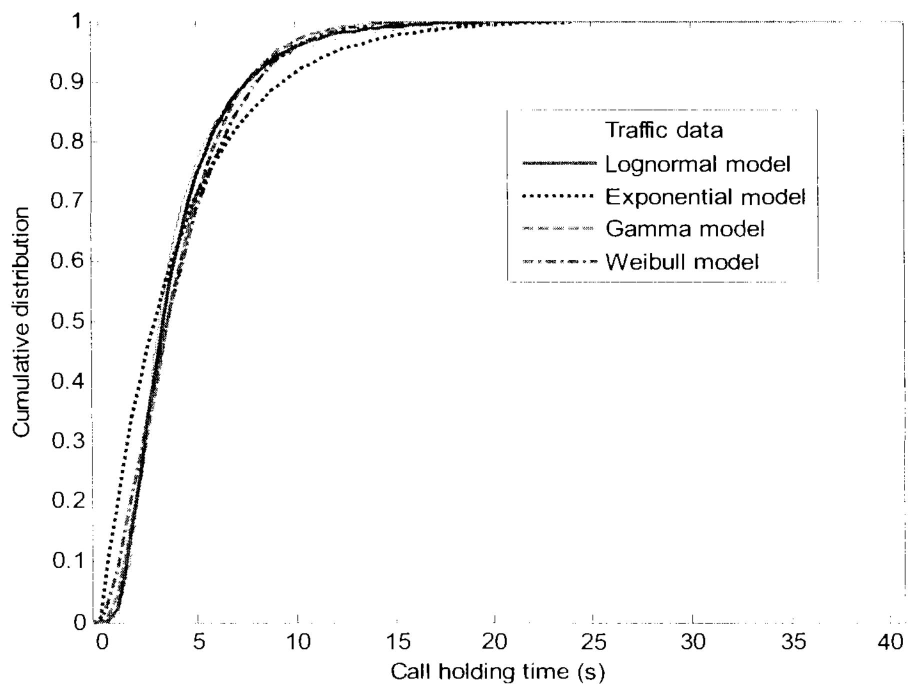

Figure 6.26: Cumulative distribution function of the call holding times and comparison with exponential, Weibull, lognormal, and gamma distributions.

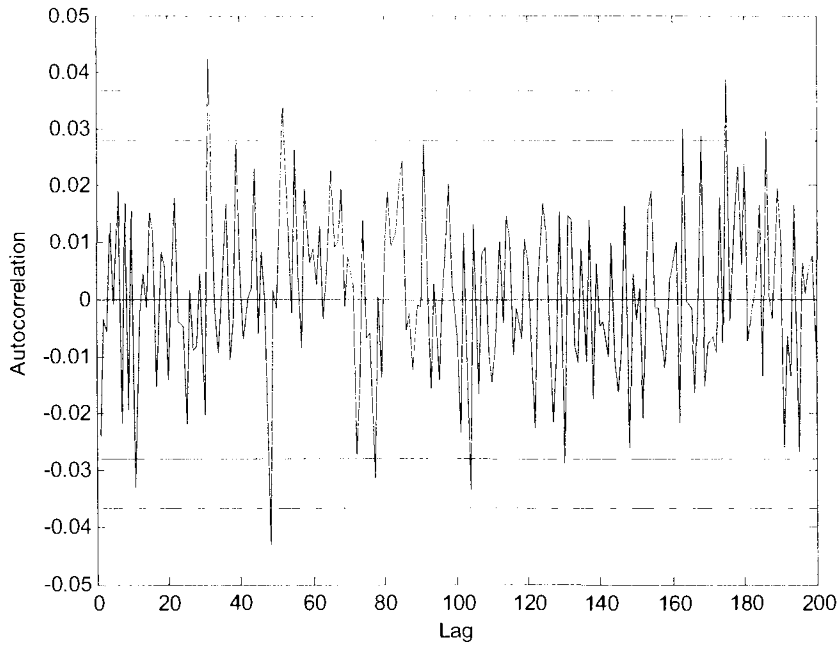

Figure 6.27: Call holding time autocorrelation plot (up to lag 200) with 95% and 99% confidence intervals.

We also investigate the long-range dependence in the call holding times. Figure 6.28 shows an example of a logscale diagram for the call holding times. The linear region has a slope of approximately zero, which implies absence of LRD. The test for time constancy of the scaling exponent α indicates that all but one estimate of H can be considered reliable. The only unreliable estimate is from the busy hour between 23:00 and 24:00 on March 26, 2003.

Table 6.9 shows the estimates of H for all hourly traces of call holding times, including the unreliable estimate (indicated by *). All estimates of H are close to 0.5, which implies that call holding times are not LRD and are uncorrelated. This is in agreement with the autocorrelation plot shown in Figure 6.27.

Table 6.9: Estimates of H for the hourly traces of call holding times.

2001		2002		2003	
Day/hour	H	Day/hour	H	Day/hour	H
02.11.2001 15:00–16:00	0.493	01.03.2002 04:00–05:00	0.490	26.03.2003 22:00–23:00	0.483
01.11.2001 00:00–01:00	0.471	01.03.2002 22:00–23:00	0.460	25.03.2003 23:00–24:00	0.483
02.11.2001 16:00–17:00	0.462	01.03.2002 23:00–24:00	0.489	26.03.2003 23:00–24:00	0.463*
01.11.2001 19:00–20:00	0.467	01.03.2002 00:00–01:00	0.508	29.03.2003 02:00–03:00	0.526
02.11.2001 20:00–21:00	0.479	02.03.2002 00:00–01:00	0.503	29.03.2003 01:00–02:00	0.466

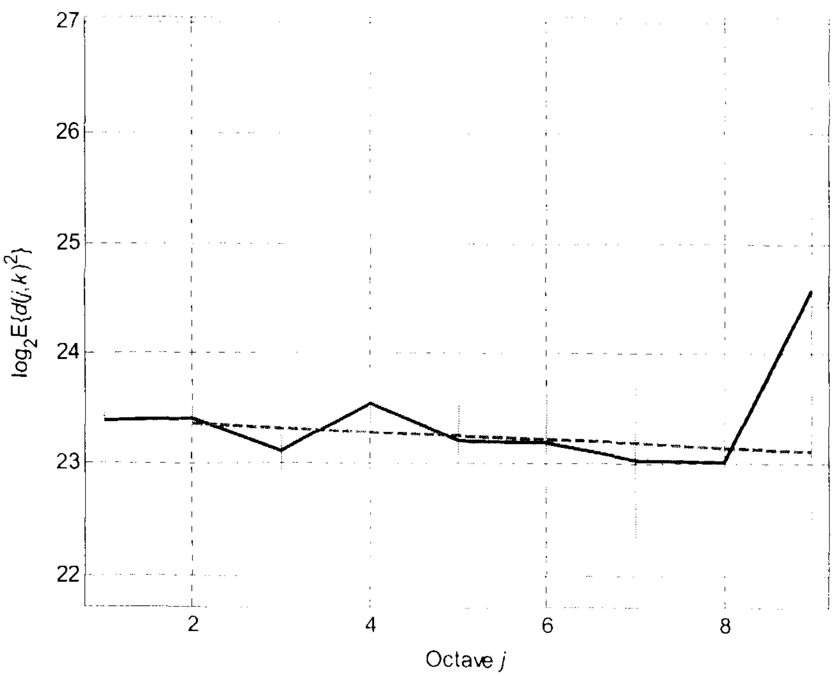

Figure 6.28: Logscale diagram of the call holding times. Dashed line is the linear regression line with slope α. Vertical lines represent the 95 % confidence intervals around the estimates of log2E{d(j,k)2}.

6.6 Discussion and Comparison of Results

Erlang B and C models, which are used to model traffic in circuit-switched networks, are based on independent exponentially distributed call inter-arrival and holding times [11]. Our analysis of the traffic from the E-Comm network indicated that neither call holding nor call inter-arrival times fit the exponential distribution. In the case of call inter-arrival times, Weibull and gamma distributions are more suitable models. In [3], only the exponential distribution was identified as the best fit. Weibull and gamma distributions are more general representations of the exponential distribution and their greater versatility makes them a better fit for the call inter-arrival times.

Nevertheless, all three distributions (exponential, Weibull, and gamma) follow closely the distribution of the traffic data, as indicated in Figure 6.21 and Figure 6.22. Lognormal distribution best fits the call holding times, which agrees with the findings from [3]. As indicated in Figure 6.25 and Figure 6.26, the distribution of call holding times significantly deviates from the exponential distribution. The parameters of the best fitting distributions for the call inter-arrival and call holding times from 2003 are shown in Table 6.10. The analytical expressions for the distributions are given in Table 6.11, where I(x) is an incomplete beta function and $\Gamma(a)$ is the gamma function.

Table 6.12 shows the average call inter-arrival and holding times for each of the busy hour traces. The average call inter-arrival times are 1.04 s, 0.86 s, and 0.84 s for the 2001, 2002, and 2003 dataset, respectively. The corresponding average call holding times are 3.91 (2001), 3.96 (2002), and 4.13 (2003).

Table 6.10: Parameters of the best fitting distributions for the call inter-arrival and call holding times for the busy hours.

Busy hour	Distribution						
	Call inter-arrival times					Call holding times	
	Exponential	Weibull		Gamma		Lognormal	
	μ	a	b	a	b	μ	σ
02.11.2001, 15:00–16:00	0.9714	0.9785	1.0175	1.0326	0.9407	1.0913	0.6910
01.11.2001, 00:00–01:00	0.9711	0.9907	1.0517	1.0818	0.8977	1.0801	07535
02.11.2001, 16:00–17:00	1.0337	1.0651	1.0826	1.1189	0.9238	1.1432	0.6803
01.11.2001, 19:00–20:00	1.0868	1.1195	1.0800	1.1295	0.9622	1.1261	0.7040
02.11.2001, 20:00–21:00	1.1170	1.1373	1.0459	1.0760	1.0381	1.0893	0.7050
01.03.2002, 04:00–05:00	0.8121	0.8313	1.0603	1.1096	0.7319	1.1746	0.6671
01.03.2002, 22:00–23:00	0.8354	0.8532	1.0542	1.0931	0.7643	1.1157	0.6565
01.03.2002, 23:00–24:00	0.8620	0.8877	1.0790	1.1308	0.7623	1.1096	0.6803
01.03.2002, 00:00–01:00	0.9082	0.9266	1.0509	1.0910	0.8325	1.1334	0.6585
02.03.2002, 00:00–01:00	0.9160	0.9400	1.0686	1.1122	0.8236	1.1610	0.6680
26.03.2003, 22:00–23:00	0.7336	0.7475	1.0475	1.0910	0.6724	1.1838	0.6553
25.03.2003, 23:00–24:00	0.8492	0.8622	1.0376	1.0762	0.7891	1.1737	0.6715
26.03.2003, 23:00–24:00	0.8546	0.8579	1.0092	1.0299	0.8292	1.1704	0.6696
29.03.2003, 02:00–03:00	0.8711	0.8918	1.0617	1.0973	0.7939	1.1929	0.6437
29.03.2003, 01:00–02:00	0.8821	0.8970	1.0425	1.0705	0.8240	1.2373	0.6214

*

Table 6.11: The best fitting distributions.

Expression	Name: parameters
$f(x) = e^{-x/\mu} / \mu$	exponential: μ
$f(x) = ba^{-b}x^{b-1}e^{-(x/a)^b} I_{(0,\infty)}(x)$	Weibull: a, b
$f(x) = x^{a-1}e^{-(x/b)} /(b^a \Gamma(a))$	gamma: a, b
$f(x) = e^{-(\ln x-\mu)^2/(2\sigma^2)} /(x\sigma\sqrt{2\pi})$	lognormal: μ, σ

It can be observed that the average call inter-arrival times decreased and call holding times increased when going from 2001 to 2002 to 2003. This implies an increase in the traffic volume and increased network utilization.

Table 6.12: Average call inter-arrival and call holding times from the hourly traces.

	2001		2002		2003	
	Day/hour	Avg. (s)	Day/hour	Avg. (s)	Day/hour	Avg. (s)
inter-arrival	02.11.2001 15:00–16:00	0.97	01.03.2002 04:00–05:00	0.81	26.03.2003 22:00–23:00	0.73
holding		3.78		4.07		4.08
inter-arrival	01.11.2001 00:00–01:00	0.97	01.03.2002 22:00–23:00	0.83	25.03.2003 23:00–24:00	0.85
holding		3.95		3.84		4.12
inter-arrival	02.11.2001 16:00–17:00	1.03	01.03.2002 23:00–24:00	0.86	26.03.2003 23:00–24:00	0.85
holding		3.99		3.88		4.04
inter-arrival	01.11.2001 19:00–20:00	1.09	01.03.2002 00:00–01:00	0.91	29.03.2003 02:00–03:00	0.87
holding		3.97		3.95		4.14
inter-arrival	02.11.2001 20:00–21:00	1.12	02.03.2002 00:00–01:00	0.91	29.03.2003 01:00–02:00	0.88
holding		3.84		4.06		4.25

Estimates of H for the traces of both call inter-arrival and call holding times from 2001, 2002, and 2003 are shown in Figure 6.29. The estimates for each year are sorted in ascending order. The horizontal axis represents the rank (position in the sorted series of estimates) and the vertical axis shows the values of the estimates. Estimates of H for the traces of call holding times from the three years are very close to each other. The difference between the largest and the smallest estimate is approximately 0.05. Estimates of H for the call inter-arrival times exhibit greater variability. The difference between the largest and the smallest estimate is approximately 0.2. Comparing the Hurst parameter estimates of the traces from the various datasets (2001, 2002, and 2003), there are no large differences or an increasing or decreasing trend across the years. Rather, our findings indicate that the Hurst parameter may be regarded as an invariant characteristic of the busy hour traffic from the E-Comm network for the datasets from 2001, 2002, and 2003.

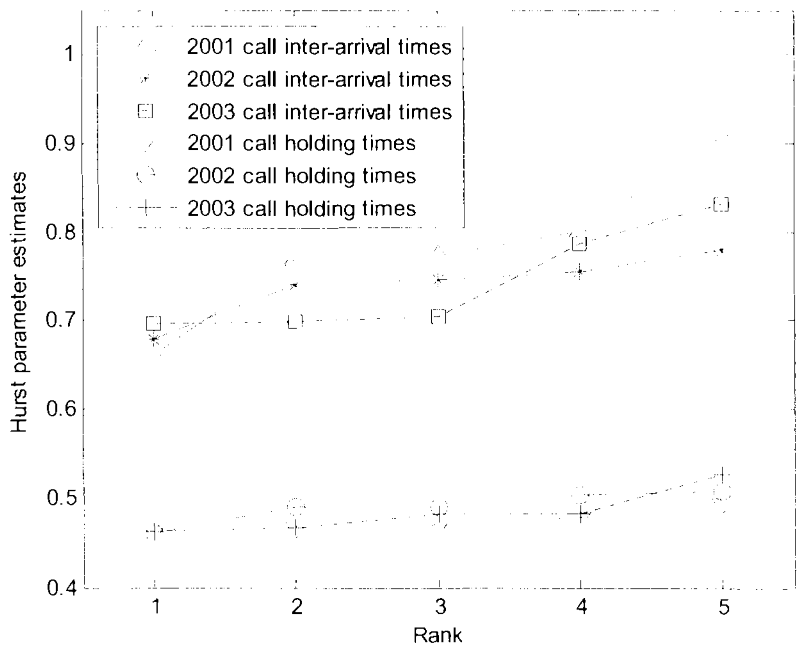

Figure 6.29: Hurst parameter estimates of the busy hour traffic traces from 2001, 2002, and 2003.

CHAPTER 7: CONCLUSION AND DISCUSSION

In this thesis, we analyzed busy hour voice traffic in a trunked, public safety wireless network from three consecutive years. We considered traffic from the cell with the largest capacity. This cell handles the majority of the calls. Our findings indicated that both Weibull and gamma distributions are suitable for modelling call inter-arrival times. Call holding times fit the lognormal distribution. Another important result of our analysis was that call inter-arrival times are long-range dependent. Therefore, traffic models that assume independent exponentially distributed call inter-arrival and holding times, such as Erlang B and C, may not produce reliable results if applied to trunked radio networks.

In addition, in this thesis, we used collected call activity data to simulate a deployed circuit-switched PSWN. We simulated network performance using the OPNET simulation tools. Simulation results show daily cycles in the network utilization. The average number of busy channels in most cells is small compared to their capacities. Nevertheless, there are busy periods of high utilization. The data model used for the trace-driven simulations only included established calls. Between February 2002 and March 2003, the number of calls increased by nearly 30%. This increase caused several cells that were underutilized in 2002 to be used near their full capacity in 2003. This simulation study may be used to address existing and future network congestion problems.

Radio networks are currently designed according to specific spectrum policies and cannot adapt to policy changes without being redesigned. This has motivated the development of software-defined radios (SDRs) and dynamic spectrum utilization [53]. The SDR radio device first "senses" the spectrum, identifies spectrum opportunities (frequency and time), and utilizes the spectrum by considering the level of interference with primary users. Hence, the management of the spectrum is placed in each radio device. The Defense Advanced Research Projects Agency (DARPA) XG (neXt Generation) communication program will enable dynamic access to temporarily unused frequencies and implementation of SDR in military and commercial areas. The XG communication protocols address the use of the spectrum, the architecture, system components, the concept of XG communications, the behaviour and organization of the XG system, and the policies for spectrum access [54].

SDRs will enable communications between public safety agencies even if they use various types of radio devices and/or reside in different cities. System level failures are of concerns for emergency communication, which require high reliability and security for their operations. Hence, the SDR technology needs to guarantee security and avoidance of system failures and interferences with adjacent frequencies.

REFERENCES

[1] N. Cackov, B. Vujičić, S. Vujičić, and Lj. Trajković, "Using network activity data to model the utilization of a trunked radio system," in *Proc. SPECTS*, San Jose, CA, July 2004, pp. 517–524.

[2] N. Cackov, J. Song, B. Vujičić, S. Vujičić, and Lj. Trajković, "Simulation and performance evaluation of a public safety wireless network: case study," *SIMULATION: Transactions of the Society for Modeling and Simulation International*, vol. 81, no. 8, pp. 571–585, Aug. 2005.

[3] D. Sharp, N. Cackov, N. Lasković, Q. Shao, and Lj. Trajković, "Analysis of public safety traffic on trunked land mobile radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 7, pp. 1197–1205, Sept. 2004.

[4] J. Song and Lj. Trajković, "Modeling and performance analysis of public safety wireless networks," in *Proc. IEEE IPCCC*, Phoenix, AZ, April 2005, pp. 567–572.

[5] L. A. Andriantiatsaholiniaina and Lj. Trajković, "Analysis of user behavior from billing records of a CDPD wireless network," in *Proc. Workshop on WLN* 2002, Tampa, FL, Nov. 2002, pp. 781–790.

[6] D. Tang and M. Baker, "Analysis of a local-area wireless network," in *Proc. ACM Mobicom 2000*, Boston, MA, Aug. 2000, pp. 1–10.

[7] D. Tang and M. Baker, "Analysis of a metropolitan-area wireless network," in *Proc. ACM Mobicom '99*, Seattle, WA, Aug.1999, pp. 13–23.

[8] E-Comm, Emergency Communication for SW British Columbia Incorporated [Online]. Available: http://www.ecomm.bc.ca.

[9] R. J. Orsulak, R. R. Seach, J. P. Camacho, and R. J. Matheson, "Land mobile spectrum planning options," *NTIA*, Washington, DC, Oct. 1995. [Online]. Available: http://www.ntia.doc.gov/osmhome/reports/slye_rpt/cover.html.

[10] K. Zdunek, "Design considerations for trunked radio systems," in *Proc. IEEE VTC*, Dallas, TX, May 1986, pp. 194–201.

[11] H. H. Hoang, R. Malhame, and C. Rosenberg, "Communication load and delay in multichannel land mobile systems for dispatch traffic: a queuing analysis," in *Proc. IEEE VTC*, Denver, CO, May 1992, pp. 773–777.

[12] P. Cohen, D. Haccoun, and H. H. Hoang, "Traffic analysis for different classes of land mobile communication systems," in *Proc. IEEE VTC*, Toronto, ON, May 1983, pp. 283-285.

[13] V. A. Bolotin, "Modeling call holding time distributions for CCS network design and performance analysis," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 3, pp. 433–438, April 1994.

[14] V. A. Bolotin, "Telephone circuit holding time distributions," in *Proc. of 14 th International Teletraffic Congress (ITC'94),* Antibes Juan-les-Pins, France, June 1994, pp. 125–134.

[15] C. Jedrzycki and V. C. M. Leung, "Probability distribution of channel holding time in cellular telephony systems," in *Proc. IEEE 46-th Vehicular Technology Conference (VTC'96),* Atlanta, GA, April 1996, vol. 1, pp. 247–251.

[16] R.A. Guerin, ''Channel occupancy time distribution in a cellular radio systems', *IEEE Transaction on Vehicular Technology*, vol.VT–35, pp 89–99, Aug. 1987.

[17] Y. Fang, I. Chlamtac, and Y. B. Lin, "Modeling PCS networks under general call holding time and cell residence time distributions," *IEEE/ACM Transaction on Networking*, vol. 5,no. 6, pp. 893–906, Dec. 1997.

[18] Y. Fang, I. Chlamtac, and Y. B. Lin, "Channel occupancy times and handoff rate for mobile computing and PCS networks," *IEEE Transaction on Computers*, vol. 47, no. 6, pp. 679–692, June 1998.

[19] Y. Fang and I. Chlamtac, "Teletraffic analysis and mobility modeling of PCS networks," *IEEE Trans. on Communications*, vol. 47, no. 7, pp. 1062–1072, July 1999.

[20] Y. Fang and I. Chlamtac, "A new mobility model and its application in the channel holding time characterization in PCS networks," in *Proc. of IEEE INFOCOM'99*, New York, NY, USA, March 1999, vol.1, pp. 20–27.

[21] Y. Fang, "Hyper-Erlang distribution model and its application in wireless mobile networks," *Wireless Networks (WINET)*, vol.7, no.3, pp. 211–219, May 2001.

[22] Y. Lin and I. Chlamtac, "Effects of Erlang call holding times on PCS call completion," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 3, pp. 815–823, May 1999.

[23] F. Barcelo , J. Jordan, and J. Paradells "Voice holding time distribution in trunked PAMR systems," *IEEE Trans. Vehicular Technology Conference (VTC'97)*, Phoenix, AZ, USA, May 1997, vol. 1, pp. 436–440, May 1997.

[24] F. Barceló and J. Jordán, "Voice holding time distribution in trunked PAMR systems," in *Proc. of IEEE Vehicular Technology Conference (VTC'97),* Phoenix, Arizona, USA, May 1997, pp. 436–440.

[25] J. Jordán and F. Barceló, "Statistical modelling of channel occupancy in PAMR systems," in *Proc. of 15 th International Teletraffic Congress (ITC'97),* Washington, D.C., USA, June 1997, pp. 1169–1178.

[26] F. Barceló and J. Jordán, "Channel holding time distribution in cellular telephony," in *Proc. of 9 th International Conference on Wireless Communication (Wireless'97)*, Alta, Canada, July 1997, vol. 1, pp. 125–134.

[27] J. Jordán and F. Barceló, "Statistical modelling of transmission holding time in PAMR systems," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM '97)*, Phoenix, AZ, Nov. 1997, vol. 1, pp. 121–125.

[28] F. Barceló and J. Jordán, "Channel holding time distribution in public telephony systems (PAMR and PCS)," *IEEE Transaction on Vehicular Technology*, vol. 49, no. 5, pp. 1615–1625, Sept. 2000.

[29] F. Barceló and J. I. Sanchez, "Probability distribution of the inter-arrival time to cellular telephony channels," in *Proc. of the 49 th Vehicular Technology Conference (VTC'99),* Houston, TX, May 1999, vol. 1, pp. 762–766.

[30] F. Barceló and S. Bueno, "Idle and inter-arrival time statistics in public access mobile radio (PAMR) systems," in *Proc. of IEEE Global Telecommunications Conference* (GLOBECOM'97), Phoenix, AZ, Nov. 1997, vol. 1, pp. 126–130.

[31] P. V. Orlik and S. S. Rappaport, "A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 788–803, June 1998.

[32] P. V. Orlik and S. S. Rappaport, "Traffic performance and mobility modeling of cellular communications with mixed platforms and highly variable mobilities," *Proceedings of the IEEE, Special Issue on Mobile Radio Centennial 86(7)*, vol. 86, no. 7, pp. 1464–1479, July 1998.

[33] M.M. Zonoozi and P Dassanayake, "User mobility modeling and characterization of mobility patterns," *IEEE Journal on Selected Areas in Communications*, vol.15, pp. 1239-1252, Sept. 1997.

[34] K.Yeo, and C.H. Jun, "Teletraffic analysis of cellular communication systems with general mobility based on hyper-Erlang characterization," in *Proc. 26 International conference on computers and industrial engineering,* Tinos, Greece, June 2002, vol42, pp. 507-520.

[35] H. Chen and Lj. Trajković, "Trunked radio systems: traffic prediction based on user clusters," in *Proc. ISWCS*, Mauritius, Sept. 2004, pp. 76–80.

[36] B. Vujičić, H.Chen, and Lj. Trajković, "Prediction of traffic in a Public Safety Network," in *Proc. ISCAS 2006*, Island of Kos, Greece, May 2006, pp.2637-2640.

[37] V. Frost and B. Melamed, "Traffic modeling for communication networks," *IEEE Communications Magazine*, vol. 33, pp. 70–80, March 1994.

[38] Diagnostic Strategies. Traffic modeling and resource allocation in call centers [Online].
Available: http://www.diagnosticstrategies.com/papers/Traffic_Modeling.pdf.

[39] G. Hess, Land-mobile radio system engineering. Norwood, MA: Artech House, 1993, pp. 249–286.

[40] B. Vujičić, N. Cackov, S. Vujičić, and Lj. Trajković, "Modeling and characterization of traffic in public safety wireless networks," in *Proc. SPECTS 2005*, Philadelphia, PA, July 20-24, 2005, pp. 214–223.

[41] Industry Canada, Channel loading guidelines [Online]. Available: http://strategis.ic.gc.ca/epic/internet/insmt-gst.nsf/en/sf08015e.html

[42] *OPNET documentation V.9.0.A*, OPNET Technologies, Inc., Bethesda, MD, 2001.

[43] WarnSim [Online]. Available: http://www.vannet.ca/warnsim.

[44] M/A-Com EDACS [Online]. Available: http://www.macom-wireless.com/products/pve/edacs.asp.

[45] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, pp. 1–15, Feb. 1994.

[46] K. Park and W. Willinger, "Self-similar network traffic: an overview," in Self-similar Network Traffic and Performance Evaluation, K. Park and W. Willinger, Eds. New York: Wiley, 2000, pp. 1–38.

[47] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch, "Wavelets for the analysis, estimation, and synthesis of scaling data," in Self-similar Network Traffic and Performance Evaluation, K. Park and W. Willinger, Eds. New York: Wiley, 2000, pp. 39–88.

[48] D. Veitch, Matlab code for estimation of scaling exponents [Online]. Available: http://www.cubinlab.ee.mu.oz.au/~darryl/secondorder_code.html.

[49] D. Veitch and P. Abry, "A statistical test for time constancy of scaling exponents," *IEEE Transaction on Signal Processing*, vol. 49, no. 10, pp. 2325–2334, Oct. 2001.

[50] R. B. D'Agostino and M. A. Stephens, Eds., Goodness-of-Fit Techniques. New York: Marcel Dekker, 1986. pp. 63–93, pp. 97–145, pp. 421–457.

[51] MSDN Library [Online]. Available: http://msdn.microsoft.com/library/default.asp?url=/library/en-us/tsqlref/ts_da-db_9xut.asp.

[52] Engineering Statistics: http://www.itl.nist.gov/div898/handbook/index.htm.

[53] J. Reed, *Introduction to Software Radio Concepts*. Upper Saddle River, NJ: Prentice Hall, 2002.

[54] DARPA, Full Spectrum Access [Online].Available: http://www.darpa.mil/ato/programs/xg/rfcs.htm.

[55] MATLAB [Online]. Available: http://www.mathworks.com/products/matlab.

[56] S-PLUS [Online]. Available: http://www.insightful.com/products/splus.

APPENDIX:
CALL_TYPE TABLE

This table represents types of calls that could be made within the E-Comm network. The group call is the most often type of call.

Call_Type	Id
Group call	0
Individual call	1
Emergency call	2
System call	3
Morse code	4
Test	5
Paging	6
Scramble	7
Group set	8
System log	9
Start emergency	10
Cancel emergency	11
N/A	100